ISCTE Business School
Instituto Universitário de Lisboa

Department of Finance

U LISBOA | UNIVERSIDADE DE LISBOA

Department of Mathematics
Faculty of Sciences

The Convolution Method for Pricing American Options under Lévy Processes

Pedro Simões Oliveira

Dissertation

Master of Science in Mathematical Finance

2014

ISCTE Business School
Instituto Universitário de Lisboa

Department of Finance

U LISBOA | UNIVERSIDADE DE LISBOA

Department of Mathematics
Faculty of Sciences

The Convolution Method for Pricing American Options under Lévy Processes

Pedro Simões Oliveira

Dissertation

Master of Science in Mathematical Finance

Supervisor: Professor Doutor João Pedro Nunes

2014

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my thesis adviser, Professor João Pedro Nunes, for his suggestions, reviewing of this manuscript and for his teachings during the course. A special thanks to Pedro Baltazar and Rui Monteiro for their help in the development of this work, and for the ever lasting discussions about the field of mathematical finance. I would also like to thank my colleagues André Sousa and Paulo Duarte for the joyful times we have shared together throughout this academic journey. To Dr. Nuno Lourenço Dimas, from InvestQuest, I would like to thank the support and motivation for finishing this thesis.

Last but not least, a special thanks to my parents, who always gave me the strength and conditions to pursue my dreams, and to my future wife Filipa, who is undoubtedly my biggest supporter.

To my fiancé Filipa.

Resumo

Um método flexível, rápido e exacto para avaliação de opções, desde as mais simples às mais complexas com provisões de exercício antecipado, é apresentado. Este método baseia-se na *Fast Fourier Transform* (FFT) e depende, naturalmente, das transformadas de Fourier. A ideia principal baseia-se em reconhecer que a fórmula usual de avaliação *neutra ao risco* pode ser calculada como uma convolução. Esta característica, é extremamente útil, dado que convoluções no domínio do tempo podem ser transformadas facilmente em multiplicações no domínio de Fourier, o que permite aplicar a FFT e beneficiar da sua capacidade computacional. Este recente método de avaliação, proposto por Lord et al. (2008), foi apelidado de método da convolução, e é aplicável a uma grande variedade de *payoffs* necessitando apenas do conhecimento da função característica do modelo. Desta forma, o método é aplicável a vários modelos afins, entre os quais está a classe de modelos exponenciais de Lévy.

O método apresentado é capaz de extender os métodos anteriores, baseados na FFT para o cálculo de opções Europeias, ao conseguir avaliar opções com provisões de exercício antecipado. Considerando-se uma opção Bermuda M vezes exercível, a complexidade gobal do método é $\mathcal{O}(MN\log(N))$, em que N é número de pontos da grelha utilizados na discretização do preço do activo subjacente. No contexto das opções Americanas, que são os contratos de opções em bolsa mais transaccionados, uma técnica eficiente, baseada na aplicação da extrapolação de Richardson aos preçoos de opções Bermudas, é apresentada.

Abstract

A flexible, fast and accurate method for pricing options, from *plain vanilla* to the more complex ones with early-exercise features, is presented. The method is based on the Fast Fourier Transform (FFT) which relies, naturally, on Fourier transformations. The key idea is to recognize that the usual *risk-neutral* valuation formula can be calculated as a convolution. This feature, is highly useful, since convolutions in the time domain can be translated easily to the Fourier domain, enabling one to apply the FFT and benefit from its computational power. This recent pricing method, proposed by Lord et al. (2008), was dubbed the convolution method, and is applicable to a wide variety of payoffs requiring only the knowledge of the characteristic function of the model. As such, the method is applicable within many regular affine models, among which is the class of exponential Lévy models.

The presented method is able to extend previous methods, based on the calculation of the FFT for pricing European options, by pricing options with early-exercise features. Considering an M-times exercisable Bermudan option, the overall complexity of the method is $\mathcal{O}(MN\log(N))$, with N grid points used to discretize the price of the underlying asset. In the context of American options, which are the most exchange traded option contracts, an highly efficient technique, based on the application of the Richardson extrapolation to the prices of Bermudan options, is presented.

Palavras Chave
Keywords

Palavras Chave

Avaliação de opções

Opções Americanas

Processos de Lévy

Convolução

Transformada de Fourier rápida

Keywords

Option pricing

American options

Lévy processes

Convolution

Fast Fourier transform

Contents

List of Figures

List of Tables

1 Introduction

Le Calcul des probabilités ne pourra sans doute jamais s'appliquer aux mouvements de la cote et la dynamique de la Bourse e sera jamais une science exacte.

Louis Bachelier, in *Théorie de la Spéculation*

This thesis presents a fast and accurate quadrature-based method for pricing options with early-exercise features, with a focus on the Fourier transform techniques and their applicability to the option pricing problem. This initial chapter starts, in Section 1.1, with a historical review of the option pricing problem that portrait the evolution of financial modelling and pricing models. Then, in Section 1.2, the framework of this work is presented, where the pricing problem is formulated and the notation is introduced. In Section 1.3, a discussion on the more sophisticated models and numerical methods is made, with emphasis on the Lévy processes and transform methods. Section 1.4 explains the approach followed in this work and what it attempts to achieve. Finally, the chapter concludes with an overview of this thesis.

1.1 Motivation

Derivatives and options, in particular, came into existence long before the opening of the Chicago Board Options Exchange (CBOE) in 1973. In ancient Babylonian and Greek times they were already cited and described by Aristotle as "a financial device which involves a principle of universal application", Bernstein (1996). In Pliska (2010) an 1885 article from *The Economist* is cited with the title "Virtues and Vices of Options", and it reports that in the Paris and in the German bourses there was a vast speculation by means of options. The article depicted options as generally great safeguards against unexpected and violent movements in prices, which favored the experienced speculators, but also remarked that they fostered a form of speculation which already flourished too abundantly. One of the most famous bubbles caused by the trading of futures contracts and options was the Dutch tulip mania of the seventeenth century. As portrayed in Bernstein (1996), the trading involved options on tulips rather than in the tulips themselves.

The world's first ever formal and standardized derivative contract was listed on the Chicago Board of Trade (CBT) when it opened in 1848; this contract was a future. The trading of standardized options contracts was only possible when the CBOE was founded in 1973 and became the first marketplace for trading listed options — see Options Exchange (2013) and Bernstein (1996). In one of the most curious coincidences, the exchanged opened just one month before the most influential article of research ever published in the field of economics or finance, the Black & Scholes (1973). This paper together with Merton (1973), for which Scholes and Merton received the Noble Prize for Economics in 1997 (Black had already died), launched the field of financial engineering.

Attempts to arrive at an option pricing formula started with Bachelier (1900). Louis Bachelier's doctoral dissertation *Théorie de la Spéculation* marks the start of the modelling of financial markets using stochastic processes and is a landmark in the history of option pricing.[1] In Merton (1997), Nobel lecture Bachelier's thesis is referred as marking the twin births of both the continuous-time mathematics of stochastic processes and the continuous-time economics of derivative pricing. There were several accomplishments in Bachelier's work, which are emphasized in Davis et al. (2011) — foreword written by the Nobel prize winner Paul Samuelson — beginning with the modelling of stocks as a Brownian motion with a drift, which means he beat Albert Einstein — who only wrote the movement equations 5 years later — in discovering what the Brownian[2] motion essentially is. He also established a connection between the Brownian motion and the Fourier's heat equation as being a diffusion process, which therefore could be applied to the calculation of diffusion probabilities. Notions of random walk or efficient markets get the important boost from Bachelier. The best forecast for tomorrow is the price today, the next price variation is independent of the previous one (no memory) and is generated by the same unknown process that drives the underlying asset. Recognized the concept of arbitrage and derived a simple formula for calculating the price of *at-the-money*[3] calls. It is quite remarkable, because he wrote long before Robert Wiener provided a coherent basis for the differential probability space, and even before Itô calculus was available. Bachelier?s work introduced, starting from scratch, much of the panoply of modern stochastic analysis, including many concepts generally associated with the names of other people working at considerably later dates. Therefore, much of the agenda for probability theory in the succeeding sixty years was concerned precisely with putting all these ideas on a rigorous footing.

Nevertheless, his work was undervalued and neglected for more than half a century — see Mandelbrot & Hudson (2010) — even by the jury of his PhD thesis, which included his mentor Henri Poincaré.[4] The study of financial markets was not yet an academic discipline, much less an appropriate topic for

[1] The Bourse at Paris was, at the time, the capital of bond trading; the depth was such that futures and options trading was developed in parallel. So Bachelier was familiar with these types of contracts as documented in Mandelbrot & Hudson (2010).

[2] A Scottish botanist, Robert Brown, studied the motion of tiny pollen grains on water and observed that it was a physical phenomenon, and received the credit for the discovery through the term "Brownian motion".

[3] Options where the strike price equals the spot price.

[4] Henri Poincaré was one of the most celebrated mathematicians of all time, that worked on every field of mathematics and beyond: probability, function theory, topology, geometry, optics and celestial mechanics.

the approval of a jury of mathematicians. Poincaré classified the theme as far from those usual treated by the candidates, although he praised its originality and suggested that the most ingenious model should have been more developed. In Courtault et al. (2000) where his role as a pioneer in both mathematical finance and probability theory is emphasized, it is argued that Poincaré had much appreciation for Bachelier's work. At that time, very few of those actively involved in financial markets had the mathematical background necessary to understand his work.

As reported in Mandelbrot & Hudson (2010) economists only begin trying to understand financial markets after the Crash of 1929. Only in the 1960s did the idea of "fair game" caught on, due to Paul Samuelson[5], and economists recognized the practical virtues of describing markets by the laws of chance and Brownian motion. Therefore, only 65 years after did the Bacherlier's work achieved renown, when the economist Paul Samuelson published an article on the pricing of warrants.[6] Bachelier's Brownian motion model for the stock price had many imperfections, such as, for example, allowing the existence of negative stock prices. In Samuelson (1965), a more realistic model was introduced, the Geometric Brownian Motion, which is still taught and studied in finance courses throughout the world. With Samuelson's work a connection was made in the 1960s between financial economics and the stochastic analysis of the day. This fact contributed for the Nobel prize he received afterwards.

In Merton (1997) Nobel lecture he reminded that before the pioneering work of the late 1950s and 1960s finance theory was little more than a collection of anecdotes, rules of thumb, and shuffling of accounting data. It was not until the end of the 1960s and early 1970s that models of finance in academia become considerable more sophisticated, involving intemporal capital asset pricing and derivative-security pricing, which employed stochastic differential and integral equations, stochastic dynamic programming, and partial differential equations. These mathematical tools were a quantum level more complex than had been used in finance before and they are still the core tools employed today. By 1975 traders on CBOE were using the Black-Scholes model to both price and hedge their option positions; it was so widely used that Texas Instruments sold a hand-held calculator specially programmed to produce Black-Scholes option prices and hedge ratios (those were the days where personal computers did not exist).

Prior to the discovery of the Black-Scholes formula, investors and speculators would have had to use heuristic methods and their projections of the future to arrive at a price for a derivative. One of the reasons for the major success of the Black-Scholes model was that it led to pricing formulas that used mostly data from observable variables. Their formulas do not require the knowledge of either investors' tastes or their beliefs about expected returns on the underlying stocks, as pointed out in Merton (1976). The previous attempts made to arrive at an option pricing formula all lacked the crucial insight of

[5]Paul Samuelson was the inventor of the option terms "American" and "European" style options.

[6]A warrant is similar to an option, but is issued and guaranteed by a company, whereas options are exchanged traded instruments.

Black, Scholes and Merton that, under certain assumptions, the risk of an option can be fully hedged by dynamically investing in the underlying asset of that option. Thus, under the assumption of no arbitrage opportunities in financial markets, the price of any option must be equal to the price of its replicating portfolio.

Considering the same replication reasoning, more exotic payoffs could be priced. An essential requirement for such a price to make sense within the option pricing model, is that the prices of actively traded instruments, such as European options, coincide with their market price. Soon, it became apparent that this was not the case in the Black-Scholes model, and that the assumption that the underlying asset follows a geometric Brownian motion with constant drift and volatility was inappropriate. The assumption was also proved to be wrong when one analyzed the implied volatility. By inverting the Black-Scholes formula with respect to the volatility for a series of options with different strikes, and with the same maturity, one should obtain, approximately the same (constant) implied volatility. However, it is a well documented empirical fact that the implied volatility is not constant as a function of strike nor as a function of time to maturity, but resembles a "smile". Hence, coming up with a pricing model which can explain both empirical phenomena has been the focus of much of the research within mathematical finance.

1.2 Framework

Options definitions and terminology

Before we introduce the problem addressed in this work, we begin by introducing some definitions and terminology associated with options, as defined in Hull (2009). In the world of derivatives, the options are the most shining stars. They are quite different from futures, forwards and swaps, mainly due to giving the holder of the option the right to buy or sell an asset, but *not* the obligation to do so. To compensate for this benefit, the holder of the option will be charged an up-front fee, the option *premium*, whereas with the other contracts it costs nothing to enter in a trade. Options are mainly exchange traded contracts for which there are two types of options. A *call* option gives one the right to buy an asset for a pre-specified price and at a specified time in the future. A *put* option gives one the right to sell an asset for a certain price and by a given date. The price specified in the contracts is known as the *strike price* or *exercise price* and the date in the contract is denominated by *expiration date* or *maturity date*. These contracts can have many different types of underlying assets such as: stocks, indices, currencies, bonds or other derivatives. In this work we will be focused on stock options. The buyer of the option is known as the *holder*, and is said to have a *long* position[7] over the underlying asset, whereas the person who sells

[7] A *long* position means that one has positive exposure to the underlying asset value, while a *short* position denotes a negative exposure to the same.

it is named as the *writer*. During the lifetime of an option the underlying asset price varies and therefore the option is also classified regarding the relation between the current asset price, S, and the contract *strike* price, K. For instance, a call where $S > K$ is referred as being *in-the-money* (ITM), if the prices are the same, $S = K$, it is *at-the-money* (ATM) and *out-of-the-money* (OTM) when $S < K$. Finally, we present the options considered in this work, which distinguish themselves by their exercise features:[8]

European options Can only be exercised at the *expiry* date.

Bermudan options Allow the exercise only at discrete points in time, not necessarily equidistant.

American options These can be exercised at any time up to the maturity of the contract.

European options are usually easier to analyze than the others, due to having only one exercise possibility. Typically, assuming a given model for the asset price, one will arrive to closed-form expressions for the prices of European options, but this will not be the case for Bermudan or American options. These, are frequently deduced or approximated from the European relatives. According to Hull (2009, p. 6), most of the options that are traded on exchanges are American.

Notation

Here some definitions and notation associated with properties of the continuous-time theory are introduced. Consider a time horizon $T \in [0, \infty]$, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a filtered probability space, where Ω is the set of all outcomes that are possible. $\mathcal{F}$ is a *sigma-algebra* containing all sets for which we want to assess, where $\mathcal{F} = \mathcal{F}_T$ and the filtration $\mathcal{F} = \{\mathcal{F}_t, t \in [0, T]\}$ satisfies the usual conditions. $\mathbb{P}$ is the physical measure which gives the probability that an event contained in the set of $\mathcal{F}$ might occur. A stochastic process $X = \{X_t, t \in [0, T]\}$ is a family of random variables defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The process is considered to be adapted to the filtration $\mathcal{F}$ if X_t is $\mathcal{F}_t$-measurable, i.e., for each t, $X_t \in \mathcal{F}_t$. Finally, the value of a stochastic process X at time t will be identified in this work by either X_t or $X(t)$.

Pricing problem

In this work we present a method for pricing American options. As mentioned before, it is impossible to obtain a closed-form expression for this type of option. Nevertheless, an approximation is obtainable if one knows how to solve a simpler problem, such as, the valuation of an European option. In Section 1.4 it will be explained how the price of an American option will be obtained from the European valuation formula. Therefore, the pricing problem is, essentially, the valuation of an European option of maturity

[8]The geographical reference in the names of options bear no relation to where they are traded.

T, with strike price K and written on the spot price of some underlying asset, S. Throughout this thesis, S_t will reference the stock-price process at time $t \geq 0$ and V_t will denote the option value at time t. The payoff of an European option at maturity, V_T is given by

$$V_T = \begin{cases} (S_T - K)^+ & \text{call option} \\ (K - S_T)^+ & \text{put option} \end{cases} \tag{1.1}$$

The price of the option at time t can be obtained using risk-neutral valuation.[9] The existance of an equivalent martingale measure $\mathbb{Q}$, also known as risk-neutral measure, where the value of every asset appreciates at the risk-free interest rate, is assumed. Under this measure, the discounted price of an asset is a martingale. Thus the premium for a call option at time t will be given by

$$\begin{aligned} C_t &= e^{-r\tau}\mathbb{E}_{\mathbb{Q}}\left[C_T|\mathcal{F}_t\right] \\ &= e^{-r\tau}\mathbb{E}_{\mathbb{Q}}\left[(S_T - K)^+|\mathcal{F}_t\right], \end{aligned} \tag{1.2}$$

where r is the risk-free interest rate, assumed to be constant, and $\tau = T - t$ is the time to maturity. The expectation in (1.2) can be calculated via numerical integration if one knows in closed-form the asset price probability density, under measure $\mathbb{Q}$. The expression in (1.2) can be further developed into

$$\begin{aligned} C_t &= e^{-r\tau}\mathbb{E}_{\mathbb{Q}}\left[(S_T - K)\mathbf{1}_{\{S_T > K\}}|\mathcal{F}_t\right] \\ &= e^{rt}\mathbb{E}_{\mathbb{Q}}\left[\frac{S_T\mathbf{1}_{\{S_T>K\}}}{e^{rT}}|\mathcal{F}_t\right] - e^{-r\tau}K\mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{\{S_T>K\}}|\mathcal{F}_t\right]. \end{aligned} \tag{1.3}$$

We can further simply (1.3), by noting that the expected value of an indicator function is a probability and that one can apply a change of numeraire from the *money-market account* to the asset price of the underlying. Let $\mathbb{Q}^S$ denote the equivalent martingale measure associated with the numeraire S_t, then we can finally write

$$C_t = S_t e^{-q\tau}\mathbb{Q}^S(S_T > K|\mathcal{F}_t) - e^{-r\tau}K\mathbb{Q}(S_T > K|\mathcal{F}_t), \tag{1.4}$$

where $\mathbb{Q}^S$ and $\mathbb{Q}$ denote the probabilities of the option ending *in-the-money* under the asset price and money-market account numeraire, respectively.

In this thesis we will only provide pricing formulas for call options, which will be denoted by C_t. The prices for put options, P_t, on the same stock with same strike and maturity, can be obtained easily using the *put-call parity*:

$$P_t = Ke^{-r\tau} + C_t - S_t e^{-q\tau}, \tag{1.5}$$

which is valid for any model — see Hull (2009, pp. 208-210) for a proof.

[9] The theory of risk-neutral valuation was first developed by Harrison & Kreps (1979).

1.3 State of the art

Asset price models

According to Schoutens (2003), much of the research within mathematical finance has been focused on alternative stochastic processes for the underlying asset, such that the prices of traded European options are more closely, if not perfectly, matched. Lévy processes were proposed in the late 1980s and early 1990s, due to being tractable and attractive models that perform significantly better that the standard Black-Scholes model. The main idea is to replace the Normal distribution of the increments by a more general one, which is able to represent the skewness and excess kurtosis present in the financial markets.

Examples of such distributions, which can take into account skewness and excess kurtosis, are the Variance Gamma (VG), the Normal Inverse Gaussian (NIG), the CGMY (named after Carr, Geman, Madan and Yor), the (Generalized) Hyperbolic Model and the Meixner distributions. Madan & Seneta (1990) and Madan et al. (1998) have proposed a Lévy process with VG distributed increments. The Hyperbolic Model was proposed by Eberlein et al. (1998). In the same year, Barndorff-Nielsen (1995) proposed the NIG Lévy process. All three previous models are special cases of the Generalized Hyperbolic Model, which was developed by Eberlein & Prause (1998). More recently, the CGMY model was introduced by Carr et al. (2002) as an extension to the VG process and the Meixner model was suggested in Schoutens (2001). Another important subclass of Lévy processes are the jump-diffusion models, which dates back to Merton (1976). A more recent model, using a different distribution for the jump sizes, was introduced by Kou (2002).

Despite the analytical tractability offered by the Lévy processes, the constraints of independence and stationarity of their increments bring some drawbacks. First, the stationarity of increments of Lévy processes leads to rigid scaling properties for marginal distributions of returns, which are not observed in empirical time series of returns, as noted by Mandelbrot & Hudson (2010). Second, under the risk neutral measure, the exponential-Lévy models are able to calibrate the implied volatility patterns for a single maturity, but fail to reproduce option prices, correctly, over a range of different maturities. Both issues are owed to the fact that exponential-Lévy models do not allow for time *inhomogeneity*, i.e., the returns, volatility and more generally the environment changes over time. As reported in Mandelbrot & Hudson (2010) or Cont & Tankov (2012), it has been observed that the estimated volatilities change stochastically over time and are clustered.

To solve this problem, extensive research has been directed into stochastic volatility models. Cox (1975) proposed the Constant Elasticity Variance (CEV) model, which attempted to introduce the leverage effect. More recently, Heston (1993) proposed another diffusion based stochastic volatility model in which the price and volatility were correlated and the later followed a *squared-root process*. However, the

diffusion based stochastic volatility models are unable to generate sufficient variability and asymmetry in short-term returns to match implied volatility skews for short maturities. A jump-diffusion stochastic volatility model was proposed by Bates (1996), which deals with this problem by adding proportional log-normal jumps to the Heston stochastic volatility model. Another possibility is to add jumps in the volatility process. Using a (positive) Lévy process to drive the volatility, it is possible to build positive, mean-reverting volatility processes with realistic dynamics without resorting to non-linear models: positive Omstein-Uhlenbeck processes, proposed as models for volatility by Barndorff-Nielsen & Shephard (2001) and Barndorff-Nielsen & Shephard (2002). These models are analytically tractable but computations become quite involved as soon as "leverage" effects are included. Finally, a different way to build models with dependence in increments is to time change a Lévy process by a positive increasing process with dependent increments. A study on this topic can be found in Carr et al. (2003).

Numerical pricing methods

Considering the risk-neutral valuation formula (1.2), one can apply several numerical techniques to calculate the price: lattice methods, Monte Carlo simulation, finite-difference methods associated with the partial (integro) differential equation (P(I)DE) and numerical integration. Possibly the most intuitive way to price options is through the use of lattice methods; the first was proposed by Cox et al. (1979) and it was a binomial tree. The use of finite-difference methods to solve option pricing problems has been around since the work of Schwartz (1977). Monte Carlo simulation in its most basic form, first suggested for use in option pricing by Boyle (1977), is probably the simplest numerical method one can implement. Lattice methods are generally less efficient and less flexible than finite-difference methods and therefore never have an advantage. As long as American features are not required and great accuracy is not neccessary, Monte Carlo is a very good method for pricing options. American features are difficult to incorporate, however, because lattice methods value the option backwards in time whereas Monte Carlo methods value forward in time and convergence to the correct solution is not exceptional. What is certainly true is that the coefficient of convergence for Monte Carlo methods is high, and getting good answers can take a long time. The main advantage Monte Carlo claims over grid-based schemes comes in its handling of multiple dimensions. Whereas for finite-difference and lattice methods the computational time increases often exponentially with the number of dimensions, for Monte Carlo methods the increase is linear. Another major advantage is that it is very flexible with regards to the choice of the distribution of returns. For problems in one or two underlying assets it is generally preferable to use finite-difference methods.

As mentioned before, the cumulative probabilities in (1.4) can be calculated provided that the density functions for the equivalent martingale measures $\mathbb{Q}$ and $\mathbb{Q}^S$ are known in closed-form. However, this is not the case for many models that, nevertheless, possess a characteristic function in closed-form.

Therefore, starting with the pioneering work of Heston (1993), much of the work has been focused on obtaining the option prices by inverting the characteristic function. The previous work was generalized by Bakshi & Madan (2000), which provided an economical interpretation of characteristic functions with respect to market completeness and that demonstrated that markets could be spanned by characteristic functions. Based on their work the cumulative probabilities in (1.4) can be found by inverting the characteristic function:

$$\mathbb{Q}(S_T > K) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\,\frac{e^{-iuk}\phi(u)}{iu}du \tag{1.6}$$

$$\mathbb{Q}^S(S_T > K) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\,\frac{e^{-iuk}\phi(u-i)}{iu\phi(-i)}du, \tag{1.7}$$

where i is the imaginary unit, k is the logarithm of the strike price K, Re denotes taking the real part of the complex number, and ϕ is the characteristic function of the log-price. This Fourier invertion process dates back to Gurland (1948) and Gil-Pelaez (1951).

Though the decomposition of an option price into probability elements is theoretically attractive, as explained by Bakshi & Madan (2000), it is numerically undesirable owing to discontinuity of the payoffs. Moreover, the pricing methods based on the calculation of this probabilities are unable to benefit from the computation power of the Fast Fourier Transform (FFT), which represents one of the most fundamental advances in scientific computing. This is due, in part, to the fact that FFT cannot be used to evaluate the integral in (1.6) since the integrand is singular at the required evaluation point $u = 0$. To tackle this issues, a numerically and very efficient methodology was introduced in Carr & Madan (1999) who pioneer the use of FFT algorithms by mapping the Fourier transform directly to call option prices via the characteristic function of an arbitrary price processes. The idea of Carr and Madan was to consider a dampened call price and to compute its Fourier inversion. Although this approach was new to the area of option pricing, the idea of damping functions on the positive real line in order to be able to find their Fourier transform is an idea that goes back to, at least, Dubner and Abate as reported in Carr & Madan (1999). Whereas Carr & Madan (1999) took the Fourier transform with respect to the strike price of the call option, Raible (2000) and Lewis (2001) used an approach which is slightly more general in that it does not require the existence of a strike in a payoff. Raible took the transform with respect to the log-forward price, Lewis used the log-spot price. Note that for all three methods, the Fourier transform of the option price can be decoupled into two parts, a payoff-dependent part, the payoff transform, and a model-dependent part, the characteristic function.

While the study of early-exercise features within the first three techniques has been plentiful, the research on the pricing via quadrature pricing techniques has not been considered until the work of Andricopoulos et al. (2003) and O'Sullivan (2005). The first paper to consider the pricing of Bermudan options using the Fourier transform inversion techniques was O'Sullivan (2005), who extended the

QUAD method of Andricopoulos et al. (2003) to allow for models where the density is not known in closed-form. Picking up on a presentation of Reiner (2001), Lord et al. (2008) proposed a method using a convolution technique, dubbed as the CONV method, which this work is about. In comparison with the QUAD method of Andricopoulos et al. (2003) and O'Sullivan's algorithm, the CONV method was able to substantially reduce the running time of the algorithm, by reducing from *quadratic* to *linearithmic*[10] the algorithm complexity in the calculation of the continuation value. This will be explained in detail in Chapter 3. The COS method was proposed in Fang & Oosterlee (2008), to solve the pricing of Bermudan options as well. In this method, Bermudan options are priced via Fourier-cosine series expansions and it has the same complexity as the CONV method. Furthermore, it is argued in Fang & Oosterlee (2008) that for certain models with characteristic functions where the rate of decay is faster than some polynomials, the convergence of the COS algorithm will be faster than in the CONV method.

1.4 Approach and main contributions

This thesis is not intended to introduced any new asset price model or numerical method. As the title suggests, it deals with the pricing of American-style options, when one considers that the asset price follows the dynamics of a Lévy process, using a numerical technique referred to as the convolution (CONV) method. The research on Lévy process has been abundant, mostly regarding the prices of European vanilla options, and the use of popular numerical methods such as Monte Carlo or finite-difference methods. Practitioners demand fast and accurate prices and sensitivities. As the financial models and option contracts used in practice are becoming increasingly complex, efficient methods have to be developed to cope with such models. The research on quadrature techniques only began in recent years, and the CONV method is one of newly ingenious created methods. The work in this thesis follows closely the paper of Lord et al. (2008) where the CONV method is proposed for the first time. Thus, this thesis goal is to:

- motivate the necessity of new fast, efficient and accurate numerical methods;
- present the fundamental Lévy definitions, properties and components while trying to avoid unnecessary technicalities and demonstrations;
- introduce the most popular Lévy models by emphasizing the rationality behind each process and the respective parameters, in order to maximize their ease of use in applications;
- provide the details of the numerical algorithms implementation for pricing under the CONV method;

[10] A *linearithmic* algorithm is the product of a *linear* and *logarithmic* algorithms. More on this in Section 3.5.

- dedicate special attention to the American-style options given that they are the most exchanged traded options contracts according to Hull (2009, p. 6).

The work in this thesis is *not* intended to:

- be a treatise on Lévy processes; for applications of Lévy processes in finance the reader should consult Papapantoleon (2008) or Schoutens (2003);
- be a study of jump-diffusion or pure jump models; Cont & Tankov (2012) provide a very complete book on jump processes;
- to provide demonstrations, for all Lévy models, of how one arrives to the European option pricing formula; for each model, references to the appropriate papers are given;
- to provide an exhaustive survey of the literature in numerical methods or Fourier transform methods.

Regarding the choice of the Lévy models to discuss in this work, the rationale was to present the most popular ones, which were also the ones studied in Lord et al. (2008). The list of models chosen is broad enough to contemplate various types of Lévy processes. We also consider one classic diffusion process, the celebrated geometric Brownian motion (GBM), which is, undoubtedly, the most studied stochastic process in the field of mathematical finance. Then, the choice was to split the Lévy processes into jump-diffusion and pure jump processes. For each type, two models were chosen with the following reasoning: one should be a classic reference, the first of its kind, and the second should be a more recent development which was intended to improve the previous models. In the class of the jump-diffusion models, the choices were to introduce the Merton (1976) model as the classic reference and the Kou (2002) model as the most recent development. Regarding the pure jump class, the variance gamma (VG) was elected as the reference on pure jump processes and for the refined model the option could only be the CGMY, which is a generalization of the variance gamma process.

The CONV method is applicable to a wide variety of payoffs and only requires the knowledge of the characteristic function of the model. As such, for every model the respective characteristic function will be presented. The accuracy of the CONV method will be evaluated first by comparing its numerical results for European-style options to the respective closed-form expressions.

1.5 Overview of this thesis

This thesis is organized as follows. Chapter 2 deals with Lévy processes. The main definitions and properties are presented and also the basic stochastic processes, which serve as the building blocks

for more complicated Lévy processes, are introduced. A discussion on jump-diffusion and pure jump processes is made and the important class of exponential Lévy processes is presented. The remaining part of Chapter 2 introduces the Lévy models considered in this thesis, namely: one diffusion model (geometric Brownian motion), two jump-diffusion models (the Merton and Kou models), and two pure jump models (the variance gamma and the CGMY).

Chapter 3 is dedicated to the method being studied in this work, the convolution (CONV) method. This method is based on a quadrature technique and relies mostly on Fourier transformations. The main idea is to reformulate the well-known risk-neutral valuation formula by recognizing that it is a convolution. The resulting convolution is dealt with numerically by using the Fast Fourier Transform (FFT). This method is applicable to a wide variety of payoffs and only requires the knowledge of the characteristic function of the model. As such, the method is applicable within many affine models. Special attention is given to the implementation details, like the problem of dealing with discontinuities. A section is also dedicated to the complexity analysis which enlightens the power of this method. The chapter terminates with the algorithm for approximating the prices of Bermudan options to the desired prices of the American ones.

In Chapter 4, the method's overall performance is evaluated in terms of speed and accuracy by pricing European, Bermudan and American-style options. For European options we compare the performance of the CONV method with other commonly used methods. Special attention is given to the pricing of American options, in particular, to the performance gains of using the suggested approximation technique. Finally, Chapter 5 concludes this thesis by outlining the main contributions of this method to the options pricing field and by discussing directions for future work.

2 Lévy Processes

Paul Lévy was a painter in the probabilistic world.

Michel Loève, in *Ann. Probability 1 (1971)*

This chapter presents and discusses the Lévy processes which will be used as the driving stochastic processes of the asset returns. We begin in Section 2.1 with the definition of a Lévy process and state its fundamental properties. Next, in Section 2.2, the differences between jump-diffusions and pure jump processes are analyzed. In Section 2.3 the class of exponential-Lévy models used for describing the asset price process are explained. In the following sections, the specific Lévy processes being considered and tested in this work are presented. Section 2.4 presents the celebrated geometric Brownian motion, which is a diffusion model. Section 2.5 discusses the first model to consider the existance of jumps, owing to Merton. In Section 2.6, the Kou's model is revealed, which is a more recent jump-diffusion model. In Section 2.7 a three parameter generalization of the Brownian motion is proposed, the so-called variance gamma model. Finally, Section 2.8 introduces another pure jump or also called *diffusion free* model named as CGMY after the authors' names.

2.1 Definitions and properties

The Lévy processes are the continuous relatives of the *random walks*, which are sums of independent and identically distributed (i.i.d) random variables. Lévy processes are at the same time simple enough to study and rich enough for application in more realistic models. Thus, stochastic models based on Lévy processes often allow for analytically or numerically tractable formulas. They are the key example of a stochastic process in continuous-time which allows for the creation of a great variety of stochastic models on top of it.

This feature is owed to the actual Lévy processes being the composition of simpler stochastic models, such as the Poisson and Wiener processes. These two processes are the fundamental examples of a Lévy process and constitute its building blocks. Before providing a formal definition of a Lévy process, the definitions for the Poisson and Wiener processes are given.

Definition 1 (Wiener process) *A stochastic process $W = \{W_t, t \geq 0\}$ is a Wiener process, also refered as a standard Brownian motion, with $W_0 = 0$, on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ if:*

1. *W has independent increments, that is, $W_t - W_s$ is independent of $\mathcal{F}_s$ for any $0 \leq s < t \leq T$.*
2. *W has stationary increments, i.e., for any $0 \leq s, t \leq T$ the distribution of $W_{t+s} - W_t$ does not depend on t.*
3. *W is stochastically continuous, i.e., $\forall t \in [0, T], \epsilon > 0 : \lim_{s \to t} \mathbb{P}(|W_t - W_s| > \epsilon) = 0$.*
4. *The increment $W_{t+s} - W_t$ is Normally distributed with mean 0 and variance $s > 0$, i.e., $W_{t+s} - W_t \sim \mathcal{N}(0, s)$.*

This is the classical example of a diffusion process, and is certainly the most studied and notorious stochastic model in quantitative finance. As mentioned before in Section 1.1, this motion was first discovered by Brown in 1928 (for which it was named after), then in Bachelier (1900) it was created as model for stock market prices, while five years later Einstein consider it as model of particles. Only in 1923 the Brownian motion was defined and constructed rigorously by Robert Wiener, for which the process is also referred. Finally, it was thanks to Samuelson (1965) that the Brownian motion was, definitely, set as the standard modelling tool in finance.

Definition 2 (Poisson process) *Consider a sequence of independent exponential[1] random variables $(\tau_i)_{i \geq 1}$ with parameter λ and $T_n = \sum_{i=1}^{n} \tau_i, n \in \mathbb{N}^+$. Then a stochastic process $N = \{N_t, t \geq 0\}$, with $N_0 = 0$, defined as*

$$N_t = \sum_{n \geq 1} \mathbf{1}_{\{t \geq T_n\}} \tag{2.1}$$

is called a Poisson process with intensity λ. This process has the following properties:

1. *N has independent increments, that is, $N_t - N_s$ is independent of $\mathcal{F}_s$ for any $0 \leq s < t \leq T$.*
2. *N has stationary increments, i.e., for any $0 \leq s, t \leq T$ the distribution of $N_{t+s} - N_t$ does not depend on t.*
3. *N is stochastically continuous, i.e., $\forall t \in [0, T], \epsilon > 0 : \lim_{s \to t} \mathbb{P}(|N_t - N_s| > \epsilon) = 0$.*
4. *The increment $N_{t+s} - N_t$ has a Poisson distribution with parameter $\lambda(t - s)$.*

When the process is characterized by a constant intensity parameter, λ, it is referred to as homogeneous. Not surprisingly, if the intensity parameter varies with time t, $\lambda(t)$, the process is classified as non-homogeneous.

The Poisson process is an increasing pure jump process, with jump sizes always equal to 1. The jumps occur at times T_i and the intervals or waiting times between jumps are exponentially distributed.

[1] *See Section A.2 of the Appendix for more details on exponential random variables.*

The French physicist and mathematician Poisson is more known for this discrete probability distribution, also named after him. That expresses the probability of a given number of events occurring in a fixed interval of time and/or space, if these events occur with a known intensity (average rate) and independently of the time since the last event. Poisson distributions are closely associated with counting activities. Observing the definitions 1-2 one can conclude easily, that only property 4 differs between the two processes, i.e., the distribution of the increments of the process is the differentiating characteristic. However, the main idea behind Lévy processes is that for a process to be analytically tractable it must not impose constraints on the distribution of the increments of the stochastic process. Next, the formal definition of a Lévy process is given:

Definition 3 (Lévy process) *A real valued and adapted stochastic process $L = \{L_t, t \geq 0\}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $L_0 = 0$ is called a Lévy process if it possesses the following properties:*

1. *L has independent increments, that is, $L_t - L_s$ is independent of $\mathcal{F}_s$ for any $0 \leq s < t \leq T$.*
2. *L has stationary increments, i.e., for any $0 \leq s, t \leq T$ the distribution of $L_{t+s} - L_t$ does not depend on t.*
3. *L is stochastically continuous, i.e., $\forall t \in [0, T], \epsilon > 0 : \lim_{s \to t} \mathbb{P}(|L_t - L_s| > \epsilon) = 0$.*

Condition 3 of the previous definition does not require that the sample paths generated by the stochastic process are continuous.[2] This can be easily verified for the Poisson process discussed previously. There is a strong relation between Lévy processes and infinitely divisible distributions, being the latter defined as:

Definition 4 (Infinite divisibility) *A probability distribution F is said to be infinitely divisible if for any integer $n \geq 2$, there exists n i.i.d. random variables $Y_1, \ldots, Y_n$ such that $Y_1 + \ldots + Y_n$ has distribution F.*

If $L = \{L_t, t \geq 0\}$ is a Lévy process, then for any $t > 0$ the distribution of L_t is infinitely divisible. This puts a restriction on the distributions that can be chosen for the increments of Lévy process for they too must be infinitely divisible. An alternative definition can be made using the characteristic function of a distribution. The definition of characteristic function associated with a random variable is given next.

Definition 5 (Characteristic function) *The characteristic function ϕ of a distribution, or equivalently, of a random variable X is the function ϕ_X defined by:*

$$\phi_X(u) = E\left[e^{iuX}\right] = \int_{-\infty}^{\infty} e^{iux} dF(x), \tag{2.2}$$

[2] Brownian motion is the only (non-deterministic) Lévy process with continuous sample paths.

where $F(x) = P(X \leq x)$ is the distribution function. A characteristic function is always continuous and verifies $\phi(0) = 1$ and $|\phi(u)| \leq 1$ for all $u \in \mathbb{R}$.

One important remark about the characteristic function of a random variable is that it completely characterizes its law, i.e., random variables with the same characteristic function are identically distributed and it is also possible to derive the moments of the random variable from ϕ. The knowledge of the characteristic function is essential to the study of Lévy processes, due to the fact that, often we do not know the distribution function of such a process in closed-form, but the characteristic function is known explicitly. Moreover, the knowing of the charateristic function plays a major role in the convolution method being studied in this work. Therefore, the alternative definition of infinitely divisible distributions using characterisitc functions is given next:

Definition 6 (Infinite divisibility with characteristic functions) *The law of a random variable X is infinitely divisible if, for all $n \in N^+$, the characteristic function $\phi_X(u)$ is also the n^{th} power of the distribution, that is, if there exists a random variable $X^{1/n}$, such that:*

$$\phi_X(u) = (\phi_{X^{1/n}}(u))^n \tag{2.3}$$

The infinitely divisible distributions together with the celebrated Lévy?Khintchine representation theorem, presented next, link the stochastic processes to distributions functions.

Theorem 1 (Lévy-Khintchine representation) *Consider a Lévy process $L = \{L_t, t \geq 0\}$ associated with a triplet (μ, σ, ν), where $\mu \in \mathbb{R}, \sigma \in \mathbb{R}_0^+$ and ν is a positive measure on $\mathbb{R}$, not necessarily finite. Then its characteristic function is given by:*

$$\phi(u) = \mathbb{E}\left[e^{iuL_t}\right] = e^{t\psi(u)}, \tag{2.4}$$

where $\psi(u)$ is the cumulant characteristic function, also known as characteristic exponent, which has the following expression

$$\psi(u) = i\mu u - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux\mathbf{1}_{\{|x|\leq 1\}}\right)\nu(dx), \tag{2.5}$$

and ν is referred to as the Lévy measure of L, which must satisfy the following properties:

1. $\nu(0) = 0$
2. $\int_{\mathbb{R}}(1 \wedge x^2)\nu(dx) < \infty$

For a proof see Cont & Tankov (2012, p. 85).

From equation (2.5) one may conclude that, in the most general case, a Lévy process consists of three independent parts or fundamental processes: a linear deterministic part where μ is called the *drift*

term, a Brownian part with a *diffusion* coefficient σ and a *pure jump* process whose dynamics is dictated by the Lévy measure $\nu(dx)$. The measure $\nu(dx)$ defines how the jumps happen, which occur according to a compound Poisson process, presented next, with intensity $\lambda = \int_{\mathbb{R}} \nu(dx)$.

Definition 7 (Compound Poisson process) *Consider a Poisson process $N = \{N_t, t \geq 0\}$ with intensity parameter λ and assume that $Y_i, i \geq 1$, is an i.i.d. (independent and identically distributed) sequence of random variables independent of N. A compound Poisson process with intensity $\lambda > 0$ and jump size distribution f is a stochastic process X_t defined as:*

$$X_t = \sum_{i=1}^{N_t} Y_i. \tag{2.6}$$

Thus, the value of the process at a given time t, X_t, is the sum of N_t random numbers following the same distribution f.

If one considers that $Y_i = 1, i \geq 1$, then the compound Process defaults to the ordinary Poisson process. In the context of jump-diffusion modelling, working with processes with a single possible jump size, like the Poisson process, is not interesting. Thus, the compound Poisson process is a generalization of the latter, where the waiting times between jumps are exponential, but the jump sizes can have an arbitrary distribution.

The Lévy measure is the main responsible for the richness of the class of Lévy processes. Considering that the Lévy process contains a Brownian motion, i.e. a *diffusion* component, and a *jump* component driven by a compound Poisson process, it is often called a *jump-diffusion* process. But it is worth nothing that not all *jump-diffusion* processes are Lévy processes.[3] The Lévy process is usually characterized by the Lévy characteristics, or Lévy triplet for short, (μ, σ, ν) which were already introduced in Theorem 1. Thus, the simplest Lévy process is a linear drift with the tripet $(\mu, 0, 0)$, and adding a *diffusion* component one gets the triplet $(\mu, \sigma, 0)$. A *pure jump* process will be identified by the triplet $(0, 0, \nu)$ and finally a (Lévy) *jump-diffusion* process will have the complete triplet (μ, σ, ν).

2.2 Jump models

The Lévy jump models can be classified into two categories: jump-diffusion or pure jump models. Another way to say it, is to referred to them as finite or infinte activity processes. The level of activity of a process is determined by the possible number of jumps in any time interval being finite or infinte. Jump-diffusion models are of finite activity and pure jump models have infinite activity processes. An excellent reference in models with jumps is Cont & Tankov (2012), which was used extensively in this thesis, and particularly in this chapter.

[3] For example, the jump-diffusion model with stochastic volatility of Bates (1996) is not a Lévy process.

Jump-diffusion models

In a jump-diffusion process its dynamics follow a diffusion process punctuated with jumps at random times, which represent rate events, such as market crashes or merger announcements. This process is simply a combination of a Brownian motion with drift and a compound Poisson process, and therefore is a process which sometimes jumps and has a continuous but random evolution between the jump times. The equation which describes a jump-diffusion process X_t is, in general, given by:

$$X_t = \mu t + \sigma W_t + \sum_{i=1}^{N_t} Y_i, \tag{2.7}$$

where μt represents the drift, σW_t the Brownian motion or Wiener process part, and finally, the last component is the compound Poisson process. As mentioned before, the jump-diffusion models have finite activity, that is, have finite jump intensity, and in this case the Lévy measure is usually rewritten as:

$$\nu(dx) = \lambda f(dx), \tag{2.8}$$

where f is the jump size distribution.

This type of models have several advantages since they have a structure that is easy to understand given that the distribution of jump sizes is known. It can also be simulated with ease using Monte Carlo for pricing path dependent options. Finally, they provide a good interpolation of the volatility smiles. A noted disadvantage is the fact that it rarely leads to closed-form densities which are necessary for some pricing methods.

Pure jump models

Pure jump models allow for an infinite number of jumps in every time interval, most of them are of small size and there is only a finite number of jumps with absolute value greater than any given positive number. It has been argued in Geman (2002) and Carr et al. (2002) that the Brownian component is not necessary with the reasoning that the jump dynamic is rich enough to describe the asset price process. In both papers it is argued that such a process gives a more realistic description of the price process at various times scales. Typically, they can be constructed via Brownian *subordination*, i.e., time changing a Lévy process with another increasing Lévy process, which gives them extra tractability. Nevertheless, they are usually harder to grasp than the jump-diffusion models, their structure is not so obvious, at least to someone used to the diffusion processes.

One could also query about finite activity models without the diffusion component, which have been proposed in Press (1967) and Cox & Ross (1976), but according to Cont & Tankov (2012) they do not lead to a realistic description of the price dynamics.

2.3 Exponential Lévy processes

In general, a Lévy process X_t can be represented in the following form:

$$X_t = \mu t + \sigma W_t + Z_t, \tag{2.9}$$

where Z_t is a jump process with (possibly) infinitely many jumps. The standard is not to model the stock price process directly as a Lévy process, but as an exponential of a Lévy process. This ensures that the log return is positive as well as with independent and stationary increments. Hence, the stock price equation can be written as

$$S_t = S_0 e^{X_t}. \tag{2.10}$$

However, for option pricing it comes in handy to use another representation that includes explicitly the mean rate of return on the stock into the definition of the exponential Lévy model:

$$S_t = S_0 e^{\mu t + X_t + \omega t}, \tag{2.11}$$

where μ is the mean rate of return on the stock and ω is called a *convexity correction* which role will become clear later. Both (2.10) and (2.11) equations are equivalent, since adding the drift term $(\mu + \omega)t$ continues to produce a Lévy process $\tilde{X}_t = (\mu + \omega)t + X_t$. Yet, the second representation is more convenient, because it allows one to specify explicitly the desired mean rate of return. Taking the expected value of the stock price S_t we have

$$\mathbb{E}\left[S_t\right] = \mathbb{E}\left[S_0 e^{\mu t + X_t + \omega t}\right] = S_0 e^{\mu t} \mathbb{E}\left[e^{X_t + \omega t}\right]. \tag{2.12}$$

Thus, we can now set the desired mean rate of return μ if we obay the following condition:

$$\mathbb{E}\left[e^{X_t + \omega t}\right] = 1 \Leftrightarrow \mathbb{E}\left[e^{X_t}\right] = e^{-\omega t}. \tag{2.13}$$

Hence, one will configure the mean rate of return μ depending on whether we want to work with the *statistical* or *risk-neutral* stock price processes. When considering an arbitrage-free market, the prices of any instrument may be calculated as the discounted expectation of its terminal payoff under the *risk-neutral* measure $\mathbb{Q}$, as exemplified before in (1.2). Therefore, in this case we will set $\mu = r - q$, and if the condition (2.13) is fulfilled, the discounted price of the stock price will be a martingale under the risk-neutral measure, that is,

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-(r-q)t} S_t\right] = S_0. \tag{2.14}$$

Recall that the characteristic function of a Lévy process X_t is given by (2.4) and if we replace $u = -i$, then we have:

$$\phi_{X_t}(-i) = \mathbb{E}\left[e^{X_t}\right] = e^{t\psi_{X_t}(-i)}. \tag{2.15}$$

Finally, joining (2.13) and (2.15) we are able to obtain the value of the *convexity correction*:

$$\omega = -\psi_{X_t}(-i) = -\frac{1}{t}\log\phi_{X_t}(-i). \tag{2.16}$$

The method that will be presented in Chapter 3 requires the knowledge of the characteristic function of the log-price, $\log S_t$, for which we are now in condition to provide a formula. This will be expressed in terms of the characteristic function of the driving Lévy process X_t. Again, applying (2.2) for $\log S_t$ we have:

$$\begin{aligned}\phi_{\log S_t}(u) &= \mathbb{E}\left[e^{iu\log S_t}\right] = \mathbb{E}\left[e^{iu(\log S_0+\mu t+X_t+\omega t)}\right] = \mathbb{E}\left[e^{iu(\log S_0+\mu t+\omega t)}e^{iuX_t}\right] \\ &= \mathbb{E}\left[e^{iu(\log S_0+\mu t+\omega t)}\right]\phi_{X_t}(u) = e^{iu\left(\log S_0+\mu t-\log\phi_{X_t}(-i)\right)}\phi_{X_t}(u).\end{aligned} \tag{2.17}$$

2.4 Geometric Brownian motion

2.4.1 Process description

The asset price model presented in this section is by far the most widely used stochastic process of modern finance. As mentioned before in Section 1.1 the Brownian motion was introduced in finance with the groundbreaking work of Bachelier (1900), which proposed to model the price of an asset S_t at the Paris Bourse as:

$$S_t = S_0 + \mu t + \sigma W_t, \tag{2.18}$$

where $S_0 > 0$ is the initial asset price. The associated stochastic differential equation is:

$$dS_t = \mu dt + \sigma dW_t. \tag{2.19}$$

This process is designated by *generalized Brownian motion* or *arithmetic Brownian motion* and is, essentially, a standard Brownian motion with a *drift*, and also allows the possibility of using different values for the *standard deviation*, by specifying a σ value different from one. Although the usage of this model was an astonishing breakthrough, it has a major flaw, namely it does not take into account that stocks have limited liability and thus they cannot be negative. To overcome this, Samuelson (1965) proposed the renowned *geometric Brownian motion* as a stochastic model for the asset price, which has an *economic* reasoning in the origin of its breeding.

The rationality behind this model is that considering an infinitesimal interval of time dt, the return of the asset price $\frac{dS_t}{S_t}$ will be decomposed in two components: a *systematic* and *random* part. Thus, assuming the that the stock's expected return is proportional to the length of the time period, one can argue that the expected increase in the asset price is given by $\mu S_t dt$ where the μ represents the expected return on the stock. Therefore, the *systematic* part of the return comes from μdt. The stochastic component of the return can be modelled in terms of variance of returns in the given interval. The stock price changes randomly, so one can assume that its return is modelled by σdW_t where dW_t represents the Gaussian noise term with variance dt which drives the stochastic process, and $\sigma > 0$ gives the power of the noise term, that is, dictates how much does the stock price fluctuate. Thereby, joining the deterministic and random components we have the following stochastic differential equation with $S_0 > 0$:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t. \tag{2.20}$$

The previous stochastic differential equation has a unique solution, which gives the asset price stochastic process:

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}. \tag{2.21}$$

This exponential function is called a *geometric Brownian motion* (GBM). One can easily see that the $\log$ of the asset return is *normally* distributed,

$$\log\left(\frac{S_t}{S_0}\right) = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t, \tag{2.22}$$

with mean $(\mu - \frac{1}{2}\sigma^2)t$ and variance $\sigma^2 t$. Hence, one can deduce that the return of asset price S_t has a *lognormal* distribution. This stochastic process is the building block of the famous Black-Scholes-Merton model for the stock price dynamics in continuous time.

Given the definition of an exponential Lévy process that we follow in this work and that was presented in Section 2.3, we redefine the GBM as the following Lévy process X_t:

$$X_t = \mu t + \sigma W_t, \tag{2.23}$$

where the drift μ has no relation to the ones in equations (2.20) and (2.21). The associated characteristic function can be retrieved with ease from (2.4) and (2.5):

$$\phi_{X_t}(u) = \exp\left(iu\mu t - \frac{\sigma^2 u^2}{2}t\right). \tag{2.24}$$

2.4.2 European option pricing formulas

We assume that the asset price follows a GBM as well as the remaining assumptions referred in Merton (1976, p. 126), and consider an European call option on a stock with price S_t, paying dividends with a yield q, with strike K and maturity T. The Black-Scholes formula for the value of the European call option at time zero is given by equation (1.4) where the *in-the-money* probabilities are given by:

$$\begin{cases} \mathbb{Q}^S(S_T > K) = \Phi(d_1), \\ \mathbb{Q}(S_T > K) = \Phi(d_2), \end{cases}$$

where Φ is the cumulative probability distribution function for the standard *normally* distributed random variable, and the parameters d_1 and d_2 can obtained from

$$d_1 = \frac{\log\left(\frac{S_0}{K}\right) + \left(r - q + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}, \tag{2.25}$$

$$d_2 = \frac{\log\left(\frac{S_0}{K}\right) + \left(r - q - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}. \tag{2.26}$$

2.5 Merton model

2.5.1 Process description

A critical assumption in the Black-Scholes model is that the underlying stock return dynamics can be described by a stochastic process with continuous sample paths. In the seminal work of Merton (1976) another option pricing formula was derived based on a more general assumption for the process of underlying stock returns, which are generated by a mixture of a continuous and jump processes. The previously asserted assumptions state that trading takes place continuously in time and that the stock price has a continuous sample path with probability one. One can easily dispute both claims, since trading in continuous-time is not possible in practice, but more importantly, empirical evidence often shows the existence of huge price changes in consecutive time instants.

Hence, Merton (1976) proposed a new model that allows for a positive probability for a stock price change of extraordinary magnitude, no matter how small the time interval between two successive observations. The remaining assumptions in the Black-Scholes model are maintained — see Merton (1976, p. 126) for a detailed explanation. The variation in the stock price is considered to be a composition of two types of changes:

normal price vibrations due to the evolution of supply and demand or other information that causes

marginal changes in the stock price. This will again be modeled as a GBM with constant variance;

abnormal changes owed to the arrival of important information regarding the firm or industry of the stock, where the information is considered to arrive at discrete points in time and that the time between this *unquiet* times is random. This part will be modeled by a jump process.

2.5.2 The Merton model

The model presentation in this subsection follows the original paper of Merton (1976) and also Matsuda (2004) which provides a clear and thorough explanation and derivation of the model. The *abnormal* price swings, i.e., jumps are modelled as events of a Poisson-driven process, where the jumps are assumed to occur independently and identically. Thus, considering an interval of time dt as small as one desires it to be, the probability that the asset price jumps is given by a Poisson process dN_t, i.e.

$$\begin{aligned} &\text{Prob (asset price jumps once in } dt) = \mathbb{P}(dN_t = 1) = \lambda dt + \mathcal{O}(dt), \\ &\text{Prob (asset price does not jump in } dt) = \mathbb{P}(dN_t = 0) = 1 - \lambda dt + \mathcal{O}(dt), \\ &\text{Prob (asset price jumps more than once in dt)} = \mathbb{P}(dN_t > 1) = \mathcal{O}(dt), \end{aligned} \tag{2.27}$$

where $\mathcal{O}(dt)$ is the order of the approximation error and λ is the mean number of jump arrivals per unit of time, commonly referred to as the intensity of the Poisson (jump) process. Upon the arrival of newly important information about the stock at time t, a jump will occur, and neglecting the continuous component, the stock price will move from S_t to S_{t+dt}. The size of jump is also assumed to be random, given by the random variable $V_t = \frac{S_{t+dt}}{S_t}$ which is assumed to be positive. The jump sizes will also be independent and identically distributed. Thus, the percentage change on the stock price owed to the jump of size V_t can be obtained from:

$$\frac{dS_t}{S_t} = \frac{S_{t+dt} - S_t}{S_t} = \frac{V_t S_t - S_t}{S_t} = V_t - 1. \tag{2.28}$$

Having explained the effect of the jump component in the returns of the stock price, one is able to incorporate this dynamics into the GBM stochastic differential equation (2.20)

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t + (V_t - 1)dN_t, \tag{2.29}$$

where μ is the drift of the process, σ is instantaneous volatility of the stock price conditional on the jump not occuring, dW_t is the standard Brownian motion, dN_t is the Poisson process with intensity λ and $V_t - 1$ accounts for the percentage change in the stock price if a jump occurs. Note that this model contains three sources of randomness (dW_t, dN_t and V_t), which are all assumed to be independent.

In equation (2.29) the σdW_t component describes the unknown return due to the *normal* price vari-

ations, and the $(V_t - 1)dN_t$ part accounts for the *abnormal* price changes. It is easy to see that, if one assumes that a jump will not occur and therefore $dN_t = 0$, the return dynamics will be identical to the Black-Scholes model, thus the expression (2.29) can also be written in the following way:

$$\frac{dS_t}{S_t} = \begin{cases} \mu dt + \sigma dW_t, & \text{if the Poisson event (jump) does not occur,} \\ \mu dt + \sigma dW_t + (V_t - 1), & \text{if the Poisson event (jump) occurs,} \end{cases} \tag{2.30}$$

where with probability one, at the most one Poisson event occurs in a time interval dt. If a jump event occurs, then the relative price change is given by $V_t - 1$. Hence, one can conclude that the resulting sample path S_t will be continuous most of the time with jumps of differing magnitudes occuring at different points in time.

The SDE in (2.29) can be solved recurring to Itô's formula for jump-diffusion processes, yielding:

$$S_t = S_0 \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t + \sum_{i=1}^{N_t} Y_i\right], \tag{2.31}$$

where $Y_i = \log(V_i)$ is the log-price jump size random variable and $\sum_{i=1}^{N_t} Y_i$ is the compound Poisson process. From (2.7) and (2.31) one can easily see that the log return of (2.31) has an associated Lévy process X_t given by:

$$X_t = \left(\mu_m - \frac{1}{2}\sigma^2\right)t + \sigma W_t + \sum_{i=1}^{N_t} Y_i. \tag{2.32}$$

Until now, no distribution has been assigned to Y (or V), only mentioned that $\{V_i\}$ is a sequence of independent identically distributed nonnegative random variables. In Merton (1976) it is further assumed a normal distribution for the jump sizes of the log price, that is:

$$Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2). \tag{2.33}$$

Therefore, the Lévy density ν_{X_t} can obtained using (2.8) and the jump size distribution from (2.33):

$$\nu_{X_t}(x) = \frac{\lambda}{\sigma_Y\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_Y)^2}{2\sigma_Y^2}\right]. \tag{2.34}$$

Having an expression for the Lévy density, the characteristic exponent can be obtained from Theorem 1:

$$\psi_{X_t}(u) = i\mu u - \frac{\sigma^2 u^2}{2} + \lambda\left(e^{i\mu_Y u - u^2\sigma_Y^2/2} - 1\right). \tag{2.35}$$

2.5.3 European option pricing formulas

The formulas for calculating the price of a European call option using the Merton jump-diffusion model is given next — for a proof see Matsuda (2004, pp. 18-21):

$$C_t = \sum_{i \geq 0} \frac{e^{-\lambda_i T}(\lambda_i T)^i}{i!} C_t^{BS}\left(T, S_t, \sigma_i, r_i\right), \tag{2.36}$$

where C_t^{BS} is the call price using the Black-Scholes formula, which was discussed in subsection 2.4.2. The remaining parameters can be retrieved using

$$\begin{cases} \lambda_i = \lambda e^{\mu_Y + \frac{1}{2}\sigma_Y^2}, \\ \sigma_i = \sqrt{\sigma^2 + \frac{i\mu_Y^2}{T}}, \\ r_i = r - \lambda\left(e^{\mu_Y + \frac{1}{2}\sigma_Y^2} - 1\right) + \frac{i\mu_Y + i\sigma_Y^2/2}{T}. \end{cases} \tag{2.37}$$

2.6 Kou model

2.6.1 Process description

A more recent jump-diffusion model was proposed in the innovative work of Kou (2002), where it was suggested a simple double exponential jump-diffusion model[4] with analytical solutions and a psychological interpretation to the same. The main motivation for this model was the fact that the models with solutions to two of the major problems in the Black-Scholes model — namely, the asymmetric leptokurtic features of asset returns and the volatility smiles — have difficulties in solving both these problems for a broad class of exotic derivatives.

The models proposed to tackle the issue of the asymmetry of the asset returns arose from very diversified fields such as: chaos theory, fractal Brownian motion, stable processes, generalized hyperbolic models and time-changed Brownian motions (see Section 2.7). These models suffered from the problem of not being able to present analytical solutions for the prices of a wide range of options, only having solutions for European-style standard options and not for more complicated and exotic options like the path-dependent options. In a parallel development, models were being proposed to incorporate the volatility smile feature present in the options markets. The most popular ones were: stochastic volatility and ARCH models, the constant elasticity model (CEV), the Merton jump-diffusion model presented previously, affine stochastic volatility and jump-diffusion models, Lévy processes and lastly implied binomial trees. Besides from possessing the same problem of not having analytical solutions

[4]The same process was proposed independently by Ramezani (1999) which approached the problem from an econometric viewpoint, with the goal of improving the fit of Merton's model to stock price data.

for path-dependent options, several models also lacked the asymmetric leptokurtic feature. For solving this issues, the Kou (2002) paper proposes a model with the following properties:

1. Explains the two empirical phenomena: asymmetric leptokurtic feature and the volatility smiles;
2. Offers analytical solutions to a wide range of option-pricing problems, including path-dependent options and interest rate derivatives;
3. Can be embedded into a rational expectations equilibrium;
4. Has a psychological interpretation.

Even though this model has such convinient properties and powerful features, this model is very simple, which makes it a very popular model. It assumes that the log-price follows a Brownian motion plus a compound Poisson process. Until here this is the same as the Merton's model presented previously, but with the jump sizes following a double exponential distribution. Furthermore, the model parameters have a financial reasoning which made them easily interpreted. The main reason for the success of the model is being able to provide analytical solutions for a variety of option-pricing problems. This is owed to the use of exponential distributions and their *memoryless* property,[5] which makes the explicit calculations possible. This is particularly important for complicated path-dependent options and interest rate derivatives. Hence, the model is able to improve the Black-Scholes model while maintaining its analytical tractability.

2.6.2 The Kou model

The only difference between this model and Merton's model presented previously in Section 2.5 is the distribution of the jump sizes Y. Kou (2002) assumes an asymmetric double exponential distribution with density given by

$$f_Y(x) = p \cdot \eta_1 e^{-\eta_1 x} \mathbf{1}_{\{x \geq 0\}} + q \cdot \eta_2 e^{\eta_2 x} \mathbf{1}_{\{x<0\}}, \qquad \eta_1 > 0,\ \eta_2 > 0, \tag{2.38}$$

with $p + q = 1$ and where p and q represent the probabilities of upward and downward jumps, respectively. Thus, the jump sizes can be viewed as

$$Y \sim \begin{cases} \xi^+, & \text{with probability } p \\ -\xi^-, & \text{with probability } q, \end{cases} \tag{2.39}$$

where ξ^+ and ξ^- are exponential random variables with means $1/\eta_1$ and $1/\eta_2$, respectively. In the case where $\eta_1 = \eta_2$ and $p = q = 1/2$ the double exponential distribution is referred to as "the first law of

[5] See the Section A.2 of the Appendix for more details on the exponential distribution and its properties.

Laplace". Given that the jump size has the following expected value

$$\mathbb{E}[Y] = \frac{p}{\eta_1} - \frac{q}{\eta_2}, \tag{2.40}$$

$$\mathbb{E}[V] = \mathbb{E}\left[e^Y\right] = p\frac{\eta_1}{\eta_1 - 1} + q\frac{\eta_2}{\eta_2 + 1}, \tag{2.41}$$

it is necessary to impose the constraint $\eta_1 > 1$ to ensure that $\mathbb{E}[V] < \infty$ and $\mathbb{E}[S_t] < \infty$. This means that the average jump size cannot exceed 100%, which is perfectly acceptable.

The Lévy density for this process can be found in the same way as before for the Merton model. One just needs to insert the double exponential distribution shown in (2.38) into the Lévy jump-diffusion density given by equation (2.8):

$$\nu_{X_t}(x) = \lambda \cdot f_Y(x) = p \cdot \lambda \cdot \eta_1 e^{-\eta_1 x}\mathbf{1}_{\{x \geq 0\}} + q \cdot \lambda \cdot \eta_2 e^{\eta_2 x}\mathbf{1}_{\{x<0\}}. \tag{2.42}$$

The characteristic exponent of the Lévy process X_t with a double exponential jump size distribution can again be found recurring to Theorem 1:

$$\psi_{X_t}(u) = i\mu u - \frac{\sigma^2 u^2}{2} + iu\lambda\left(\frac{p}{\eta_1 - iu} - \frac{q}{\eta_2 + iu}\right). \tag{2.43}$$

2.6.3 European option pricing formulas

The formulas for the *in-the-money* probabilities are given next — for a proof see Kou (2002, pp. 1097-1100):

$$\begin{cases} \mathbb{Q}^S(S_T > K) = \Upsilon\left(r + \frac{1}{2}\sigma^2 - \lambda\xi, \sigma, \tilde{\lambda}, \tilde{p}, \tilde{\eta_1}, \tilde{\eta_2}; \log(K/S(0)), T\right) \\ \\ \mathbb{Q}(S_T > K) = \Upsilon\left(r - \frac{1}{2}\sigma^2 - \lambda\xi, \sigma, \lambda, p, \eta_1, \eta_2; \log(K/S(0)), T\right). \end{cases} \tag{2.44}$$

where we have

$$\tilde{p} = \frac{p}{1+\xi} \cdot \frac{\eta_1}{\eta_1 - 1}, \qquad \tilde{\eta_1} = \eta_1 - 1, \tag{2.45}$$

$$\tilde{\eta_2} = \eta_2 + 1, \quad \tilde{\lambda}(\xi + 1), \quad \xi = \frac{p\eta_1}{\eta_1 - 1} + \frac{q\eta_2}{\eta_2 + 1} - 1. \tag{2.46}$$

2.7 Variance gamma

2.7.1 Process description

The variance gamma (VG) stochastic process is a three parameter generalization of the Brownian motion as a model for the dynamics of the log-price. This process can be simply described as evaluating a

Brownian motion with constant drift and volatility at a *random time change* that follows a gamma process. Thus, this process is usually referred as a *time changed* Brownian motion with constant drift. The Brownian motion will be measured conditional on the realization of a random time, which will be driven by a gamma density. The variance gamma was originally proposed by Madan & Seneta (1990) by considering a time change of a Brownian motion without a drift given by a gamma process. Their paper describes the mathematical details of the process and is currently termed the symmetric variance gamma process. In the following year, Madan & Milne (1991) considered an equilibrium option pricing model using this symmetric VG process and provided an economic motivation for its use. The most cited paper regarding this process is Madan et al. (1998), which we follow in this section, and that extends the previous work of Madan & Milne (1991) by providing closed-form solutions for the return densities and prices for European options on stocks, as well as empirical support that validates its use.

Besides the volatility of the Brownian motion, this model has two additional parameters that allow one to control the kurtosis or the skewness of the distribution. For the kurtosis parameter we are able to increase, symmetrically, both left and right tail probabilities of the return distribution. Regarding the skewness, one is able to create asymmetric tails, being this part an extension to the Madan & Milne (1991) work. One important feature about this model, is the fact that it can be configured to behave as the Black-Scholes model.

The introduction of this model was at the time an important advance in the option pricing theory, due to the fact that it proposed a model with no continuous martingale component. Most of the studies considered diffusion processes that have a martingale component with continuous sample paths. The VG process is pure jump process that has high levels of activity, i.e., the price changes with intense frequency. It belongs to the class of infinite activity models, which allow for an infinite number of jumps in any time interval. Unlike the Black-Scholes model, this process has finite variation, therefore it can be written has the difference of two increasing processes. The first process reflects the price increases and the second accounts for the price decreases. Both processes are also gamma processes. Hence, the VG process may be expressed in two different forms, namely:

1. Time-changed (subordinated) Brownian motion, where the subordinator is a gamma process.
2. Difference of two gamma processes.

The next two subsections will expose the process in these two different ways.

2.7.2 VG as Brownian motion with a drift

Consider a Brownian motion with constant drift θ and volatility σ given by

$$b(t;\theta,\sigma) = \theta t + \sigma W(t), \tag{2.47}$$

where $W(t)$ is the standard Brownian motion. Also, consider a gamma process $\gamma(t;\mu,v)$ with independent increments and with mean rate μ and variance rate v, which is described in detail in Section A.3 of the Appendix.

The three parameter VG process $X(t;\sigma,\theta,v)$ is defined in terms of the Brownian motion with drift $b(t;\theta,\sigma)$ and the gamma process with unit mean rate $\gamma(t;1,v)$ as

$$X(t;\sigma,\theta,v) = b(\gamma(t;1,v);\theta,\sigma), \tag{2.48}$$

where now it is clear that the process $X(t)$ is a Brownian motion with a drift evaluated at a gamma time change. As mentioned before, the model has two additional parameters, besides the volatility σ of the Brownian motion, namely: v the variance rate of the gamma time change and θ the drift of the Brownian motion. One will be able to control the skewness of the distribution via θ and the kurtosis using v. This will become more clear after presenting the density function for the VG process. As explained in Madan et al. (1998), this can be expressed, conditional on the realization of a gamma time change, as a normal density function. The unconditional density may then be obtained by integrating out the gamma density for every increment. Therefore, the density for $X(t)$ is given by

$$f_{X(t)}(x) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi g}} \exp\left(-\frac{(x-\theta g)^2}{2\sigma^2 g}\right) \frac{g^{\frac{t}{v}-1}\exp(-\frac{g}{v})}{v^{\frac{t}{v}}\Gamma(\frac{t}{v})} dg. \tag{2.49}$$

In a similiar way, the expression for the characteristic function for the VG process can be obtained by first conditioning on the gamma time and then using the characterisitc function of the normal to obtain the conditional characteristic function. Hence, we can then employ the density of the gamma increment to integrate out the gamma time with respect to its density, giving the following expression:

$$\phi_{X(t)}(u) = \left(\frac{1}{1 - i\theta v u + (\frac{\sigma^2 v}{2})u^2}\right)^{\frac{t}{v}}. \tag{2.50}$$

2.7.3 VG as a difference of gamma processes

Madan et al. (1998) show that the VG process can also be expressed as the difference of two independent and increasing gamma processes, with the following expression:

$$X(t;\sigma,\theta,v) = \gamma_p(t;\mu_p,v_p) - \gamma_n(t;\mu_n,v_n), \tag{2.51}$$

where the $\gamma_p(.)$ refers to the process with the price increases, while the process $\gamma_n(.)$ is related with the price decreases. The relation between the parameters of the gamma processes in (2.51) and the

paremeters of the VG process in (2.48) is given by

$$\mu_p = \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{v}} + \frac{\theta}{2} \tag{2.52}$$

$$\mu_n = \mu_p - \theta \tag{2.53}$$

$$v_p = \mu_p^2 v \tag{2.54}$$

$$v_n = \mu_n^2 v. \tag{2.55}$$

Therefore, using the expression for the Lévy measure of a gamma function — see Section A.3 of the Appendix — the Lévy measure for the VG process can be obtained as

$$\nu_{X(t)}(x) = \begin{cases} \dfrac{\mu_n^2}{v_n} \dfrac{\exp\left(-\frac{\mu_n}{v_n}|x|\right)}{|x|} & \text{for} \quad x < 0, \\ \\ \dfrac{\mu_p^2}{v_p} \dfrac{\exp\left(-\frac{\mu_p}{v_p}|x|\right)}{|x|} & \text{for} \quad x > 0. \end{cases} \tag{2.56}$$

From equation (2.56) we can observe that the VG process inherits the infinite arrival rate of price jumps from the gamma process. In order to assess the role of the original VG process parameters, the Lévy measure is rewritten next, as a function of these terms:

$$\nu_{X(t)}(x) = \frac{\exp(\theta x/\sigma^2)}{v|x|} \exp\left(-\frac{\sqrt{\frac{2}{v} + \frac{\theta^2}{\sigma^2}}}{\sigma}|x|\right). \tag{2.57}$$

When we set $\theta = 0$ in (2.57) the Lévy density becomes symmetric about zero and the process has no skewness. This yields the symmetric VG process described in Madan & Seneta (1990) and Madan & Milne (1991). When $\theta < 0$, we observe from (2.57) that the negative values of x receive a higher weight than the positive ones, thus for negative values of θ the density is negatively skewed, whereas for positive values the density is positively skewed. Regarding v, for large values of v the exponential decay rate is slower around zero, and therefore it raises the probability of large jumps which lead to higher tail probabilities and excess kurtosis.

2.7.4 European option pricing formulas

For the VG process the closed-form density function is available and the respective expression was shown in equation (2.49). Thus under the risk-neutral setting one can calculate the call option price using (1.4). Conditional on g, the process $X(t)$ is normally distributed and the option formula will be similar to the Black-Scholes model. Hence, the European option price can be calculated by integrating that conditional Black-Scholes formula over g with respect to the gamma density. The formulas for the

in-the-money probabilities are given next — for a proof see Madan et al. (1998, pp. 98-102):[6]

$$\begin{cases} \mathbb{Q}^S(S_T > K) = \Psi\left(d\sqrt{\dfrac{1-c_1}{v}}, (\alpha+s)\sqrt{\dfrac{v}{1-c_1}}, \dfrac{t}{v}\right), \\ \\ \mathbb{Q}(S_T > K) = \Psi\left(d\sqrt{\dfrac{1-c_2}{v}}, \alpha\sqrt{\dfrac{v}{1-c_2}}, \dfrac{t}{v}\right). \end{cases} \tag{2.58}$$

The function Ψ is defined in terms of a modified Bessel function and degenerate hypergeometric functions of two variables — see Madan et al. (1998, pp. 98-102). The remaining parameters can be retrieved using

$$d = \frac{1}{s}\left[\log\left(\frac{S_0}{K}\right) + rt + \frac{t}{v}\log\left(\frac{1-c_1}{1-c_2}\right)\right], \tag{2.59}$$

with $\alpha = \xi s$ and ξ, s, c_1 and c_2 are defined as

$$\xi = -\frac{\theta}{\sigma^2}, \qquad s = \frac{\sigma}{\sqrt{1+\left(\dfrac{\theta}{\sigma}\right)^2 \dfrac{v}{2}}}, \qquad c_1 = \frac{v(\alpha+s)^2}{2}, \qquad c_2 = \frac{v\alpha^2}{2}. \tag{2.60}$$

2.8 CGMY

2.8.1 Process description

Pure jump processes have been used as models for stock returns since Press (1967) and Cox & Ross (1976), which studied processes with finite activity. A decade later the research was focused on processes with infinite activy. Two examples of that are the Madan & Seneta (1990) VG process, which was described in detail in Section 2.7, and the Hyperbolic Model introduced in Eberlein et al. (1998). The rationale for the usage of jump-diffusion processes was that the diffusion component captured the small moves whereas the jump component addressed the large moves. But since infinite activity pure jump processes have the ability to capture both frequent small moves and rare large moves, it raises the question as to whether the diffusion component is necessary.

Having this in mind, a new continuous-time model, dubbed the CGMY (after the authors names Carr, Geman, Madan and Yor) for the asset returns was proposed in Carr et al. (2002). The CGMY model is general enough to incorporate both diffusion and jumps of both (in)finite activity and (in)finite variance. The parameters of the process allow for finite or infinite variation of the jump part. The model has closed-form expressions for the characteristic function of the log-price, but not for the return

[6] In Madan et al. (1998) the closed-form expression in page 88 contains a small mistake, the term αs should be replaced by α, the correct expression is written in the page 99 of the Appendix.

densities. The knowledge of the characteristic function enables the description of many of the processes properties. Their research concluded that the index returns tend to be pure jump processes of infinite activity and finite variation, while for equity returns the diffusion component is virtually insignificant.

2.8.2 The CGMY model

The CGMY process is a generalization of the variance gamma process with an additional parameter to allow for finite or infinite activity and finite or infinite variation. The generalization is best viewed by presenting the Lévy density with parameters C, G, M and Y:

$$\nu_{X(t)}(x) = \begin{cases} C\dfrac{\exp(-G|x|)}{|x|^{1+Y}} & \text{for} \quad x < 0 \\ \\ C\dfrac{\exp(-M|x|)}{|x|^{1+Y}} & \text{for} \quad x > 0, \end{cases} \tag{2.61}$$

where[7] $C \geq 0$, $G \geq 0$, $M \geq 0$ and $Y < 2$. Hence, the CGMY process is an infinitely divisible process of independent increments denoted by $X(t; C, G, M, Y)$ and is defined by the Lévy density given by (2.61). If we set $Y = 0$ the process defaults to the VG process with the following parameters:

$$C = \frac{1}{v}, \tag{2.62}$$

$$G = \frac{1}{\mu_n}, \tag{2.63}$$

$$M = \frac{1}{\mu_p}. \tag{2.64}$$

As should be expected by the discussion in the VG process, the parameters are extremely important in capturing the several aspects of the stochastic process. The parameter C represents a measure of the overall level of activity, which is not surprising given its relation to the variance of the gamma time change displayed in (2.62). As mentioned in Section 2.7.3, Madan et al. (1998) show that the parameter C provided control over the kurtosis of the distribution of $X(t)$, when the Lévy measure is symmetric. This happens when $G = M$ these parameters control the rate of exponential decay of left, G, and right,M tails of the Lévy density. Thus, when $G \neq M$ it leads to skewed distributions. If $G < M$ the distribution is negatively skewed due to the left tail being heavier than the right one. This situation is consistent with the risk-neutral distributions that are usually implied from the options market prices. From (2.61) it is also possible to conclude that the exponential factor in the numerator of the Lévy density is the reason for all the moments of the process being finite.

The parameter Y is useful in characterizing the "fine structure of the stochastic process" as referred

[7]The condition $Y < 2$ is required in order for the Lévy density to integrate x^2 in the neighborhood of 0.

in Carr et al. (2002, p. 311). That is, it describes the properties and behavior of the Lévy density, specially near zero. Essentially, using the value of the parameter Y, we are able to classify the process in terms of:

Monoticity of Lévy density A completely monotone density will require that the frequency of large jumps is smaller than the frequency of small jumps.

Finiteness of activity Processes of infinite activity suit as an approximation for high liquid markets with large activity.

Finiteness of variation Processes of finite variation are more flexible in the sense that they do not require any parametric restriction when changing from the statistical to risk-neutral measures.

Although for this model the probability density function is not available in closed form, the characteristic function is readily available thanks to the Theorem 1. Thus, the characteristic function for the CGMY process is given by:

$$\phi_{X(t)}(u) = \exp\{tC\Gamma(-Y)[(M-iu)^Y - M^Y + (G+iu)^Y - G^Y]\}. \tag{2.65}$$

2.8.3 European option pricing formulas

For this process there is no probability density function available in closed-form. As such, the risk-neutral valuation formula (1.4) cannot be used directly to calculate the option price. Thereby, we are restricted to use a pricing method based on the knowledge of the characteristic function. One alternative could be to go allong the path of Bakshi & Madan (2000) and calculate the *in-the-money* probabilities using (1.6) and (1.7). Yet, in this work we decided to follow the same approach of Carr et al. (2002) where the Carr & Madan (1999) method was used for the computation of the model's option price.

3 The Convolution Method

FFT is the most important algorithm of the 20^{th} century.

Gilbert Strang, Professor of Mathematics at M.I.T.

The present chapter describes in detail the convolution (CONV) method for pricing options with early-exercise features. Section 3.1 introduces notation regarding early-exercise options and presents an algorithm for pricing Bermudan-style options. In Section 3.2 the CONV method is explained in detail and the respective algorithm presented. Section 3.3 examines the discretization of the convolution and the applied quadrature techniques. Next, in Section 3.4, guidelines on suitable grid definitions are provided, which are able to cope with points of discontinuity. Section 3.5 enlightens the importance of the FFT usage in the context of option pricing. Lastly, in Section 3.6 an efficient algorithm for approximating the prices of a series of Bermudan-style options to the price of an American-style option is unraveled. This chapter draws mostly from Lord et al. (2008) where this method was originally proposed.

3.1 Pricing problem approach

One key advantage of the CONV method, when compared with other transform based methods such as Carr & Madan (1999) or Lewis (2001), is its ability for pricing early-exercise options. Most of the options with early-exercise features are Bermudan or American. American options can be exercised at any time prior to option's maturity, while Bermudan options can only be exercised at pre-specified dates, typically with a certain frequency. Next, an algorithm for pricing Bermudan options via backward induction will be presented, and in Section 3.6 an extrapolation technique for pricing American options through the prices of the Bermudan options will be described.

Before turning to the backward induction algorithm, some notation is introduced. We define the set of exercise dates as $\mathcal{T} = \{t_1, \ldots, t_M\}$ where $0 = t_0 \leq t_1 < \cdots < t_M$. The time between the exercise dates does not need to be constant, but for ease of exposure the exercise dates are considered equally spaced, i.e., for all m we have $\Delta t = t_{m+1} - t_m$. The holder of the option can exercise the it at any time $t_m \in \mathcal{T}$ obtaining the exercise payoff $E(t_m, S(t_m))$, which will depend on the strike K and option type in the

following way:

$$E(t_m, S(t_m)) = \begin{cases} (S(t_m) - K)^+ & \text{call option} \\ (K - S(t_m))^+ & \text{put option} \end{cases} \tag{3.1}$$

The Bermudan valuation can be achieved using the backward induction described next in Algorithm 1.

Algorithm 1 Pricing a Bermudan option via backward induction

1: $V(t_M, S(t_M)) = E(t_M, S(t_M))$
2: For $m = M - 1, \ldots, 1$
3: $\quad C(t_m, S(t_m)) = e^{r\Delta t}\mathbb{E}[V(t_{m+1}, S(t_{m+1})|\mathcal{F}_{t_m}]$
4: $\quad V(t_m, S(t_m)) = \max\{C(t_m), E(t_m, S(t_m))\}$
5: end
6: $V(t_0, S(t_0)) = C(t_0, S(t_0))$

S denotes the asset on which the contract is based, V is the value of the option immediately prior to the exercise opportunity, and C is the continuation value. In Step 3 the continuation value is calculated by the application of the *risk-neutral* valuation formula, where the expectation is taken with respect to all the information available at time t_m and under the *risk-neutral* measure $\mathbb{Q}$. The interest rate r was assumed to be deterministic to simplify the exposure. The continuation value can be written as

$$C(t_m, S(t_m)) = e^{-r\Delta t} \int_{-\infty}^{\infty} V(t_{m+1}, y) f(y|S(t_m)) dy, \tag{3.2}$$

where $f(y|S(t_m))$ represents the probability density describing the transition from $S(t_m)$ at t_m to y at t_{m+1}. So, equation (3.2) requires the knowledge of the probability density function $f(y|S(t_m))$ in closed-form, which for some models is available — for example, the Black-Scholes model or the Merton's jump-diffusion model — but for many the transition density is difficult to get, whilst the characteristic function is easily obtained. So the main motivation of the transform methods, such as the Convolution, is to calculate the continuation value using the characteristic function as opposed to the transition density.

3.2 Method details

In this section, the convolution method is described in detail and a new formula for the continuation value will be obtained. First, we define x and y as monotone functions of the asset price, where x and y represent the log-spot price at times t_m and t_{m+1}, respectively. This method's main premise is the assumption that the conditional probability density in (3.2), now defined as $f(y|x)$, only depends on x and y via their difference:

$$f(y|x) = f(y - x). \tag{3.3}$$

This assumption is certainly valid when considering Lévy processes, since one of its defining properties is that its increments are independent of each other as presented in Chapter 2. So replacing (3.3) in (3.2)

and applying the following change of variables.

$$z = y - x, \tag{3.4}$$

the continuation value can be expressed as

$$C(t_m, x) = e^{-r\Delta t} \int_{-\infty}^{\infty} V(t_{m+1}, x + z) f(z) dz, \tag{3.5}$$

which can be easily seen as a convolution after the following definitions have been presented:

Definition 8 (Convolution) *The convolution of two functions $x(t)$ and $h(t)$ is defined as*

$$x(t) * h(t) \equiv \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau, \tag{3.6}$$

where "$*$" denotes the convolution operator.

Closely related with the definition of convolution is the definition of Cross-Correlation:

Definition 9 (Cross-Correlation) *The cross-correlation of two functions $x(t)$ and $h(t)$ is defined by*

$$x(t) \star h(t) \equiv \bar{x}(-t) * h(t) \equiv \int_{-\infty}^{\infty} x(\tau) h(t + \tau) d\tau \tag{3.7}$$

where "$\star$" denotes the cross-correlation operator.

Analyzing (3.5) one can conclude that the integral part can be viewed as a cross-correlation of the option value at time t_{m+1} and the transition density $f(z)$, or, in the same way, as a convolution of $V(t_{m+1})$ and the conjugate of $f(z)$. Thus, (3.5) can also be written as

$$C(t_m, x) = e^{-r\Delta t} \big[f(x) \star V(t_{m+1}, x) \big] = e^{-r\Delta t} \big[\bar{f}(-x) * V(t_{m+1}, x) \big]. \tag{3.8}$$

This insight, that the continuation value can be seen as a convolution is extremely important, because of the following Theorem:

Theorem 2 (Time-Convolution) *If $h(t)$ has the Fourier transform $H(f)$ and $x(t)$ has the Fourier transform $X(f)$, then $h(t) * x(t)$ has the Fourier transform $H(f)X(f)$.*

This means that if the Fourier transforms for both functions exist, one can replace the complicated calculation of the convolution in the time domain, by a multiplication in the frequency domain. The Fourier transform is defined in the next Theorem:

Theorem 3 (Fourier Transform) *The Fourier integral is defined by the expression*

$$\hat{h}(u) \equiv \mathcal{F}[h(t)](u) \equiv H(u) = \int_{-\infty}^{\infty} e^{iut} h(t) dt, \tag{3.9}$$

if the integral exists for every value of the parameter u, then equation (3.9) defines $H(u)$, the Fourier transform of $h(t)$.

The practice of calculating the convolution using Fourier transforms is very common in many fields of science, because the multiplication in the frequency domain is usually simpler. It is always possible to invert the resulting Fourier function in the frequency domain to the respective time domain, the next Theorem gives the formula for accomplishing this inversion:

Theorem 4 (Inverse Fourier Transform) *The inverse Fourier transform is defined as*

$$h(t) \equiv \mathcal{F}^{-1}[\hat{h}(u)](t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iut} \hat{h}(u) du, \tag{3.10}$$

The inversion transformation of (3.10) allows the determination of a function of time from its Fourier transform.

But before, we take the Fourier transform of (3.5), one must dampen the continuation value with $e^{\alpha x}$ to ensure the existence of the Fourier transform. The difference between this approach and the Carr & Madan (1999) method is that in this method the transform will be taken with respect to the log-spot price instead of the log-strike price, which was also considered for European option prices in Lewis (2001). Thereby, applying the damping factor we obtain

$$c(t_m, x) = e^{-r\Delta t} e^{\alpha x} \int_{-\infty}^{\infty} V(t_{m+1}, x+z) f(z) dz, \tag{3.11}$$

where the notation convention introduces that small letters indicate damped quantities, e.g., $c(t_m, x) = e^{\alpha x} C(t_m, x)$. Taking the Fourier transform of $c(t_m, x)$ one obtains

$$\begin{aligned}
\mathcal{F}[c(t_m, x)](u) &= e^{-r\Delta t} \int_{-\infty}^{\infty} e^{\alpha x} e^{iux} \int_{-\infty}^{\infty} V(t_{m+1}, x+z) f(z) dz dx \\
&= e^{-r\Delta t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{iu(x+z)} e^{-iuz} e^{\alpha x} V(t_{m+1}, x+z) f(z) dz dx.
\end{aligned}$$

Using the fact that $v(t_m, x+z) = e^{\alpha(x+z)} V(t_{m+1}, x+z)$, then

$$\begin{aligned}
\mathcal{F}[c(t_m, x)](u) &= e^{-r\Delta t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{iu(x+z)} e^{-iuz} e^{\alpha x} e^{-\alpha(x+z)} v(t_{m+1}, x+z) f(z) dz dx \\
&= e^{-r\Delta t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{iu(x+z)} v(t_{m+1}, x+z) e^{-iz(u-i\alpha)} f(z) dz dx.
\end{aligned} \tag{3.12}$$

Now remembering from (3.4) that $x = y - z$ and changing the order of integration, we obtain

$$\begin{aligned}
\mathcal{F}[c(t_m, x)](u) &= e^{-r\Delta t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{iuy} v(t_{m+1}, y) dy e^{-iz(u-i\alpha)} f(z) dz \\
&= e^{-r\Delta t} \int_{-\infty}^{\infty} e^{iuy} v(t_{m+1}, y) dy \int_{-\infty}^{\infty} e^{-iz(u-i\alpha)} f(z) dz \\
&= e^{-r\Delta t} \int_{-\infty}^{\infty} e^{iuy} e^{\alpha y} v(t_{m+1}, y) dy \phi(-(u - i\alpha)) \\
&= e^{-r\Delta t} \mathcal{F}[e^{\alpha y} V(t_{m+1}, y)](u) \phi(-(u - i\alpha))
\end{aligned} \tag{3.13}$$

Now, we are ready to calculate the continuation value and, therefore, to apply the backward induction Algorithm 1. After calculating the Fourier transform of the damped continuation value in (3.13), one is able to recover the undampend continuation value $C(t_m, x)$ by applying the Fourier inversion and undpaming it. This procedure for the convolution algorithm is outlined next in Algorithm 2.

Algorithm 2 The Convolution algorithm for Bermudan options

1: Set the option value at maturity $V(t_M, x) = E(t_M, x)$ for all log-spot prices x
2: Initialize the exercise payoff $E(t_0, x) = 0$ for all log-spot prices x
3: For $m = M - 1, \ldots, 1$
4: Dampen $V(t_{m+1}, x)$ with $e^{\alpha x}$ and take its Fourier transform
5: Compute $\mathcal{F}[c(t_m, x)](u)$ using (3.13)
6: Calculate $C(t_m, x)$ by applying Fourier inversion to (3.13) and undamping
7: Calculate the option value $V(t_m, x) = \max\{E(t_m, x), C(t_m, x)\}$
8: end
9: Calculate the current option value $V(t_0, x) = C(t_0, x)$

3.3 Discretization

This section deals with the discretization of the integrals associated with the convolution method. The essence of the CONV method is the calculation of the convolution:

$$c(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \hat{v}(u) \phi(-(u - i\alpha)) du, \tag{3.14}$$

where $\hat{v}(u)$ is the Fourier transform of v:

$$\hat{v}(u) = \int_{-\infty}^{\infty} e^{iuy} v(y) dy. \tag{3.15}$$

To be able to use the FFT we have to switch to logarithmic coordinates. For this reason the state variables x and y will represent $\log S(t_m)$ and $\log S(t_{m+1})$, up to a constant shift. The integrals in (3.14) and (3.15) will be approximated by discrete sums, so that the FFT algorithm can be employed in their computation. First, it is necessary to define the uniform grids associated with the variables in the integrals to be

discretized:

$$u_j = u_0 + j\Delta u, \qquad x_j = x_0 + j\Delta x, \qquad y_j = y_0 + j\Delta y, \quad j = 0, \ldots, N-1. \tag{3.16}$$

The x and y grids have the same mesh size, $\Delta x = \Delta y$, and the Nyquist relation must be satisfied to avoid *aliasing*, that is

$$\Delta u \cdot \Delta y = \frac{2\pi}{N}. \tag{3.17}$$

Each integral in (3.14) and (3.15) will be approximated using a different numerical integration rule. For the Fourier transform $\hat{v}$ one will use the general Newton-Cotes rule that stems for the following definition:

Definition 10 (Newton-Cotes Rule) *The Newton-Cotes formula for n points allows one to integrate a function $f(x)$ over some interval $[a, b]$, by dividing it into n equal parts such that $f_n = f(x_n)$ and $h = \dfrac{b-a}{n}$ is the step between each grid point, with $x_i = x_0 + nh$. The closed Newton-Cotes formula with n points is given by:*

$$\int_a^b f(x)dx \approx h \sum_{i=0}^{n} w_i f(x_i), \tag{3.18}$$

where w_i represents the weight used for each grid point x_i.

Thus, applying the approximation (3.18) to (3.15) we obtain:

$$\hat{v}(u) \approx \Delta y \sum_{n=0}^{N-1} w_n e^{iu_j y_n} v(y_n). \tag{3.19}$$

The weights w_n on the Newton-Cotes rule will be chosen as:

$$w_0 = \frac{1}{2}, \quad w_{N-1} = \frac{1}{2}, \quad \text{and} \quad w_n = 1 \quad \text{for } n = 1, \ldots, N-2. \tag{3.20}$$

Concerning the integral in (3.14) the left-rectangle integration rule will be applied, and its definition is presented next:

Definition 11 (Left-Rectangle Rule) *The rectangle method computes an approximation to an integral, by finding the area of a collection of rectangles whose heights are determined by the values of the function. The rectangles are then drawn so that either their left or right corners, or the middle of their top line lies on the graph of the function, with bases running along the x-axis. The approximation to the integral is then calculated by adding up the areas (base multiplied by height) of the rectangles, giving the formula, for left-rectangle:*

$$\int_a^b f(x)dx \approx h \sum_{n=0}^{N-1} f(x_n) \tag{3.21}$$

where $h = \dfrac{b-a}{N}$ *and* $x_n = a + nh$.

Hence, using the approximation in (3.21) in the (3.14) yields

$$c(x_p) \approx \frac{\Delta u}{2\pi} \sum_{j=0}^{N-1} e^{-iu_j x} \phi(-(u_j - i\alpha))\hat{v}(u_j) \tag{3.22}$$

Inserting (3.19) into (3.22), one obtains the following approximation for the continuation value with $p = 0, \ldots, N-1$:

$$c(x_p) \approx \frac{\Delta u \Delta y}{2\pi} \sum_{j=0}^{N-1} e^{-iu_j x_p} \phi(-(u_j - i\alpha)) \sum_{n=0}^{N-1} w_n e^{iu_j y_n} v(y_n). \tag{3.23}$$

Replacing the grid definitions of (3.16) into (3.23), one obtains:

$$c(x_p) \approx \frac{\Delta u \Delta y}{2\pi} \sum_{j=0}^{N-1} e^{-i(u_0 + j\Delta u)(x_0 + p\Delta x)} \phi(-(u_0 + j\Delta u - i\alpha)) \sum_{n=0}^{N-1} w_n e^{i(u_0 + j\Delta u)(y_0 + n\Delta y)} v(y_n), \tag{3.24}$$

which, after some algebra, yields:

$$c(x_p) \approx \frac{e^{-iu_0(x_0 + p\Delta y)}}{2\pi} \Delta u \sum_{j=0}^{N-1} e^{-ijp\frac{2\pi}{N}} e^{ij(y_0 - x_0)\Delta u} \phi(-(u_j - i\alpha))\hat{v}(u_j), \tag{3.25}$$

where the Fourier transform $\hat{v}(u)$ is approximated by

$$\hat{v}(u_j) \approx e^{iu_0 y_0} \Delta y \sum_{n=0}^{N-1} e^{ijn\frac{2\pi}{N}} e^{inu_0 \Delta y} w_n v(y_n). \tag{3.26}$$

For reasons to be described later in Section 3.4, the initial value for the Fourier integration grid will be set as $u_0 = -\frac{N}{2}\Delta u$, which implies that

$$e^{inu_0 \Delta y} = \underbrace{e^{-in\frac{N}{2}\Delta u \Delta y} = e^{-in\frac{N}{2}\frac{2\pi}{N}}}_{\text{by (3.17)}} = e^{-in\pi} = (e^{-i\pi})^n = (-1)^n.$$

Therefore, equations (3.25) and (3.26) can be recombined and further simplified into:

$$c(x_p) \approx e^{iu_0(y_0 - x_0)}(-1)^p \sum_{j=0}^{N-1} e^{-ijp\frac{2\pi}{N}} \underbrace{e^{-ij(y_0 - x_0)\Delta u} \phi(-(u_j - i\alpha))}_{x_j} \frac{1}{N} \sum_{n=0}^{N-1} e^{ijn\frac{2\pi}{N}} \underbrace{(-1)^n w_n v(y_n)}_{x_n}. \tag{3.27}$$

Now using, the following definition of the Discrete Fourier Transform (DFT):

Definition 12 (Discrete Fourier Transform) *Considering a discrete time series* x_j *with* $j = 0, \ldots, N-1$, *the*

discrete Fourier transform (DFT) is given by

$$\mathcal{D}_n\{x_j\} = \sum_{j=0}^{N-1} e^{-ijn\frac{2\pi}{N}} x_j, \tag{3.28}$$

and the its inverse transform

Definition 13 (Inverse Discrete Fourier Transform) *Considering a discrete time series x_n with $n = 0, \ldots, N-1$, the inverse discrete Fourier transform (IDFT) is given by*

$$\mathcal{D}_j^{-1}\{x_n\} = \frac{1}{N}\sum_{j=0}^{N-1} e^{ijn\frac{2\pi}{N}} x_n, \tag{3.29}$$

in conjunction with (3.27) finally leads to:[1]

$$c(x_p) \approx e^{iu_0(y_0-x_0)}(-1)^p \mathcal{D}_p\big\{e^{ij(y_0-x_0)\Delta u}\phi(-(u_j - i\alpha))\mathcal{D}_j^{-1}\{(-1)^n w_n v(y_n)\}\big\}. \tag{3.30}$$

3.4 Dealing with discontinuities

This section focuses on the construction of a suitable grid that allows the CONV algorithm to achieve a smooth convergence. This property is most welcome, since an extrapolation technique will be used in Section 3.6 to price American-style options. A Bermudan option has several discontinuities, one for each exercise opportunity, thus if the integration domain is not splitted such that we are only integrating continuous functions, the order of convergence will be affected. To illustrate this feature, two grid constructions will be described in this section, where the first does not deal with discontinuities and the second does, named as Discretizations I and II, respectively. Before investigating how to handle discontinuities in the CONV method, the grid choice for the basic Discretization I is presented.

Until now, nothing has been said about the grid initial points u_0, x_0 and y_0 or regarding the integration domain length. The latter can be identified as L and will be chosen such that most of the mass of the underlying density is inside the interval $[-\frac{L}{2}, \frac{L}{2}]$. Therefore, the grid step can be defined as $\Delta x = \frac{L}{N}$ and the initial grid points come naturally as $u_0 = y_0 = x_0 = -\frac{L}{2}$. By equating the grids x and y, the grid definitions of (3.16) can be restated as:

$$u_j = -\frac{L}{2} + j\Delta u, \qquad x_j = y_j = -\frac{L}{2} + j\Delta y, \quad j = 0, \ldots, N-1, \tag{3.31}$$

[1] In Lord et al. (2008) the definitions of the DFT and IDFT have the minus sign switched and therefore the final equation of $c(x_p)$ has the DFT exchanged with the IDFT.

but given that $L = \Delta x \cdot N$ the grid can be further simplified as[2]

$$u_j = \left(j - \frac{N}{2}\right)\Delta u, \qquad x_j = y_j = \left(j - \frac{N}{2}\right)\Delta x, \quad j = 0, \ldots, N-1. \tag{3.32}$$

As referred in Section 3.3, x and y represent, up to constant shift, $\log S(t_m)$ and $\log S(t_{m+1})$, respectively. If one considers $x = \log S(t_m) - \log S(t_0)$ and $y = \log S(t_{m+1}) - \log S(t_0)$, then x and y represent log-returns and this will be referred to as the Discretization I. A convenient property of this discretization is that the spot price always lies on the grid, given that when $j = \frac{N}{2}$ we have that $x_j = 0$ and $\log \frac{S(t_0)}{S(t_0)} = 0$.

Although more sophisticated approximations can be made regarding the choice of L, in this work we use the rule of thumb devised in O'Sullivan (2005) and applied in Lord et al. (2008), which chooses L as a multiple of the standard deviation of the log-spot price.

$$L = \delta \cdot \sqrt{Var[\log S(t_m)]} = \delta \cdot \sqrt{-\left.\frac{\partial^2 \phi(t_m, u)}{\partial u^2}\right|_{u=0} + \left(\left.\frac{\partial \phi(t_m, u)}{\partial u}\right|_{u=0}\right)^2}, \tag{3.33}$$

where $\phi(t_m, u)$ is the characteristic function of $\log S(t_m)$ conditional upon $\log S(t_0)$ and δ is a proportionality constant. The choice of L involces a trade-off, given that $\Delta y = L/N$, the Nyquist relation implies $\Delta u = 2\pi/L$ and therefore $[u_0, u_{N-1}] = [-N\pi/L, (N-2)\pi/L]$. Thus, although setting larger values of L implies smaller truncation errors, it also makes the range of the grid in the Fourier domain to be smaller, which will create a larger error initially.

Now, a grid construction is described that allows us to place a discontinuity on the grid, which is named as Discretization II. Consider that at time t_m one wishes to place the discontinuity d_m on the grid, then the grid should be shifted by a small amount in the following way:

$$x_j = \epsilon_x + \left(j - \frac{N}{2}\right)\Delta y, \qquad y_j = \epsilon_y + \left(j - \frac{N}{2}\right)\Delta y, \tag{3.34}$$

where $\epsilon_x = d_m - \lceil dm/\Delta x \rceil \cdot \Delta x$ and ϵ_y is chosen in a similar way. Even though European options only have one time step, this feature is also useful, since by choosing $\epsilon_y = \log K/S(t_0)$ and $\epsilon_x = 0$ one ensures that the discontinuity of the payoff lies on the y-grid as well as the spot price is on the x-grid. For Bermudan options this process is more complicated given that one only knows the location of one discontinuity, which is located at the final exercise time t_M. During each time step of Algorithm 1, after approximating the option value V, all that is known is that the discontinuity is inside an interval of width Δx, for example, $[x_l, x_{l+1}]$. Thus a solution with reduced costs is obtained by using a linear interpolation to locate the discontinuity d_m:

$$d_m \approx \frac{x_{l+1}(C(t_m, x_l) - E(t_m, x_l)) - x_l(C(t_m, x_{l+1}) - E(t_m, x_{l+1}))}{(C(t_m, x_l) - E(t_m, x_l)) - (C(t_m, x_{l+1}) - E(t_m, x_{l+1}))}. \tag{3.35}$$

[2]In Lord et al. (2008) the term $\frac{N}{2}$ was, in our opinion, mistakenly written as $\frac{L}{2}$.

Thereby, the grid can be shifted in such a way that d_m lies on it and the continuation and exercise values can be recalculated. It is important to note that the inner IDFT of (3.30) does not need to be recalculated, the only term that is affected is the outer DFT. This operations result in an accurate and efficient algorithm for valuing Bermudan options. The pseudo-code can be found next in Algorithm 3.

Algorithm 3 The details of the algorithm for pricing Bermudan options

Ensure that the strike K is placed on the grid by setting $\epsilon_y = \log K/S(t_0)$
For $m = M-1, \ldots, 1$
    Equate x-grid to the y-grid at time t_{m+1}.
    Compute $C(t_m, x)$ through (3.30)
    Locate the exercise boundary $[x_l, x_{l+1}]$
    Approximate d_m using (3.35)
    Set $\epsilon_x = d_m$ and recompute the x-grid
    Recalculate $C(t_m, x)$ using the new x-grid
    Calculate the option value $V(t_m, x) = \max\{E(t_m, x), C(t_m, x)\}$
    Set the y-grid at t_m to be equal to the x-grid at t_m
end
Place the initial spot price on the grid by setting $\epsilon_x = 0$
Calculate $V(t_0, x) = C(t_0, x)$ using (3.30)

3.5 Complexity considerations

In this section we follow Sedgewick & Wayne (2011) and discuss the computational speed of the CONV method. An important question that arises when one is solving complex problems or trying to process large augments of data is: *How long will my program take?* The answers depend on many factors such as the computer being used, the dimension of the data being processed and the actual program doing the job, which implements some algorithm. Thus, trying to estimate the running time of a program seems a daunting task. Nevertheless, one can get useful answers to the given problem using simple and straightforward methods. These methods use mathematical models with the objective of modeling the computational costs which can thereafter be validated through experimental studies. The running time of a program can be determined by two primary factors: (i) the cost of executing each statement; and (ii) the frequency of execution of each statement. The former factor is a property of the computer, operating system, compiler and so on, whereas the latter is a property of the program (algorithm) and the input data. Thus, one can focus on determining the frequency of execution of each statement as a way of measuring the *order of growth* of the running time of a program.

Frequency analysis can create complicated and lengthy mathematical expressions, hence the *tilde notation* ($\sim$) is often used for discharging low-order terms which complicate formulas and represent a negligible contribution to cost function. For example if one considers a problem of input size N and the function determined via frequency analysis is $N^3/6 - N^2/2 + N/3$, then the tilde approximation is given by $\sim N^3/6$ which has an associated order of growth of N^3 (constant factors are absorbed given

that they are machine dependent). The functions commonly encountered in studying order of growth of program's running time are shown in Table 3.1. A common measure for the order of growth is the *big-Oh*

Table 3.1: Commonly encountered order of growth functions.

Description	Order of growth	Operation
constant	1	statement
logarithmic	$\log N$	divide in half
linear	N	loop
linearithmic	$N \log N$	divide and conquer
quadratic	N^2	double loop
cubic	N^3	triple loop
exponential	2^N	exhaustive search

notation which is useful in providing asymptotic upper bounds on the performance of algorithms, its formal definition is given next.

Definition 14 (Big-Oh notation) *The function $f(N)$ is $\mathcal{O}(g(N))$ if there exist constants c and N_0 such that $|f(N)| \leq c|g(N)|$ for all $N > N_0$.*

Now we are able to access the complexity of the CONV method for pricing an M-times exercisable Bermudan option, with N grid points used to discretize the price of the underlying asset. Analyzing Algorithm 3 using the previous definitions, one is able to split the frequency operations in two components: (i) the loop with M operations, one for each exercise opportunity; and (ii) the calculation of the continuation value using (3.30) for a grid of size N. The latter operation is, essentially, the computation of the discrete Fourier transform (DFT) with N points, which before the invention of the Fast Fourier Transform (FFT) had a complexity of $\mathcal{O}(N^2)$, that yielded an overall complexity of $\mathcal{O}(MN^2)$. With the advent of the FFT, the complexity of the calculation of the DFT was reduced to $\mathcal{O}(N \log N)$ and thereby setting the overall method complexity in $\mathcal{O}(MN \log N)$.

The QUAD method for pricing Bermudan-style options, of Andricopoulos et al. (2003) and extended by O'Sullivan (2005) that was briefly discussed in Section 1.3, has an overall complexity of $\mathcal{O}(MN^2)$. Therefore, the usage of the FFT in the CONV method provided an algorithm with a *linearithmic* order of growth for the loop inner calculations which is much better than the previous *quadratic* algorithms. The importance of this feature is illustrated in Figure 3.1 where the running time of the algorithms is plotted against the problem size.

3.6 American options

The algorithms described previously in Sections 3.1-3.4 enable the pricing of Bermudan options with M exercise opportunities, while in this section we describe how to price American options using the prices for the respective Bermudan options. Currently, there are two fundamental approaches to solve

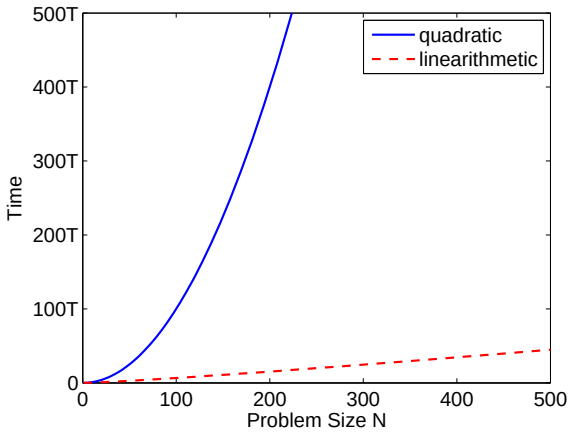


Figure 3.1: Orders of growth for *linearithmic* and *quadratic* algorithms.

this problem. The first, which is the easiest and most commonly used, approximates the American option price by considering a Bermudan option with many exercise opportunities. The second, uses a technique known as the Richardson extrapolation as a way to efficiently approximate the value of an American option using the values of Bermudan option prices. The latter approach was introduced for the first time in the innovative work of Geske & Johnson (1984). However, their method possesses two important problems, as mentioned in Chang et al. (2007), namely:

Non-uniform convergence This problem arises when the Bermudan option with n exercise opportunities has a value less than one with m exercise opportunities, where $m < n$.

Difficulty in determining the accuracy It is not possible to specify a given level of accuracy and to determine how many options and/or how many exercise points should be considered.

So, in order to solve these issues some changes to the Geske & Johnson (1984) methodology were proposed in Chang et al. (2007). First, it was suggested the usage of geometric exercise points instead of arithmetic exercise points, which allowed them to derive a modified Geske & Johnson (1984) formula which uses only the prices of Bermudan options with the uniform convergence property. Secondly, they devised a technique dubbed as the repeated Richardson extrapolation which allows one to determine the smallest number of exercise points, n, that can solely be used in an option value approximation for a specified level of accuracy.

Now, a description of the repeated Richardson extrapolation is given following Chang et al. (2007). Consider that the price of a Bermudan option with exercise times equally spaced is given by $V_{\Delta t}(t_0, S(t_0))$ where Δt denotes the time between two consecutive exercise dates. Let $A(t_0, S(t_0))$ designate the price of American option, then

$$\lim_{\Delta t \to 0} V_{\Delta t}(t_0, S(t_0)) = A(t_0, S(t_0)), \tag{3.36}$$

If one knows the complete expansion of the truncation error about the function $V_{\Delta t}$, then using this methodology we are able to expand $V_{\Delta t}$ to the price of the American option with respect to Δt. Consider that $V_{\Delta t}$ has the following expression:

$$V_{\Delta t}(t_0, S(t_0)) = A(t_0, S(t_0)) + \sum_{j=1}^{k} a_j \Delta t^{\gamma_j} + O(\Delta t^{\gamma_{k+1}}), \tag{3.37}$$

where the exponents γ_j are known with $\gamma_1 < \gamma_2 < \ldots < \gamma_j$, but the constants a_j for $j = 1, \ldots, k$ are unknown. The term $O(\Delta t^{\gamma_{k+1}})$ represents the sum of terms of order $\Delta t^{\gamma_{k+1}}$ and higher. The extrapolation method consists, essentially, in writing equation (3.37) for several different values of Δt, where for each equation a smaller value of Δt is used and consequently the number of exercise opportunities increases. The Bermudan option values, $V_{\Delta t_i}$, are known, so joining all previous equations one is able to eliminate all unknown constants a_j. According to Chang et al. (2007), using the previous construction and if we consider $\gamma_j = \gamma j, j = 1, \ldots, k$, the following algorithm can be elaborated:

Algorithm 4 The repeated Richardson extrapolation

1: For $m = 1, \ldots, k-1$
2: For $i = 1, \ldots, \max\{k, I\} - m$
3: Calculate exercise points $\Delta t_i = f(\Delta t, i)$ using (3.38)
4: Calculate Bermudan option prices $A_{i,0} = V_{\Delta t_i}$
5: Obtain the m times repeated extrapolation using (3.39)
6: end
7: end

In Step 3 the exercise points are calculated for the given number of Bermudan options to use, i, using either arithmetic or geometric exercise points. This is translated in the given equation (3.38), where Δt represents the maturity T of the option.

$$\Delta t_i = f(\Delta t, i) = \begin{cases} \dfrac{\Delta t}{i} & \text{arithmetic exercise points,} \\ \\ \dfrac{\Delta t}{2^{i-1}} & \text{geometric exercise points.} \end{cases} \tag{3.38}$$

As mentioned before, the original Geske & Johnson (1984) formula used arithmetic exercise points, which has the problem of, in some cases, leading to non-uniform convergence. To overcome this issue, Chang et al. (2007) proposed the use of geometric exercise points which will be the ones used in the experiments performed in this work. The value of $A_{i,m}$ is an approximation of the American option value $A(t_0, S(t_0))$ obtained from an m times repeated Richardson extrapolation using step sizes of $\Delta t_i, \Delta t_{i+1}, \ldots, \Delta t_{i+m}$ for $1 \leq m \leq k-1$, and is given by:

$$A_{i,m} = A_{i+1,m-1} + \frac{A_{i+1,m-1} - A_{i,m-1}}{\left(\frac{\Delta t_i}{\Delta t_{i+m}}\right)^{\gamma} - 1}. \tag{3.39}$$

4 Numerical Results

No amount of experimentation can ever prove me right; a single experiment can prove me wrong.

Albert Einstein, Nobel Prize in Physics.

Having presented all components necessary for pricing options, we turn our attention to the convolution (CONV) method performance. This chapter presents several experiments that show the accuracy and speed of the CONV method for pricing European, Bermudan and American-style options. First, in Section 4.1, the setup of the numerical experiments is described. Then, in Section 4.2, we evaluate the sensitivity of two important parameters, α and L. Next, in Section 4.3, we evaluate the CONV method in terms of speed, accuracy and convergence using the models presented in Chapter 2. Section 4.4 shows how the CONV method performs when there are multiple exercise opportunities. Finally, in Section 4.5, we test the CONV method for pricing American-style options, and compare the performance of using one Bermudan option with many exercise opportunities to the case of using a few Bermudan options with less exercise dates in combination with an extrapolation technique introduced in Section 3.6.

4.1 Experimental setup

The programming was done using MATLAB and all tests were performed on a Intel Core i5 processor with 2.4GHz and 8GB RAM. In the experiments in which the running time of a method is measured, the resulting CPU time (in milliseconds) is obtained after averaging the times of 1000 experiments. The method's flexibility is demonstrated by showing results for five asset price processes: GBM, VG, CGMY as well as Merton's and Kou's jump diffusion processes which will hereafter be identified as MJD and KJD, respectively. Table A.1 of the Appendix A contains the parameter sets used in the tests performed throughout this chapter. The pricing problems being considered are of European, Bermudan and American-style. The tests consider different option types such as *calls* and *puts*, and also shorter or longer maturities.

The error will be presented as an absolute value and is calculated as $V(t_0, S(t_0)) - V_{ref}(t_0, S(t_0))$, where the reference value $V_{ref}(t_0, S(t_0))$ is obtained by one of three ways: (i) using a closed-form solution; (ii) via another numerical scheme; or (iii) using the CONV method with many (2^{20}) grid points.

Before each result is presented, the source of the reference value will always be clarified. The other two numerical schemes considered were chosen since they to only requiring the knowledge of the characteristic function of the log-spot price as the CONV method. Moreover, they are also the common choices when pricing European-stype options. The first one is the inversion of the characteristic function by Bakshi & Madan (2000) hereafter identified as BM, which calculates option prices using equations (1.6) and (1.7). The second method is the Carr & Madan (1999) which will be referred to as CM.[1]

The rate of convergence of the algorithm will be calculated using a different definition from the commonly used for evaluating the convergence of an algorithm. Typically, one uses the Q or R-rates as one of the key performance measures of an iterative optimization algorithm — see Nocedal & Wright (2000, pp. 28-30). However, in a discretization method the important parameter for the convergence speed is not the number of iterations, but the number of used grid points. Hence, we follow an alternative definition for the rate of convergence.

Definition 15 (Convergence speed for discretization methods) *A sequence x_n with $n \in \mathbb{N}^+$ is said to converge to L with order p, if there exists a constant C such that*

$$|x_n - L| \leq C \cdot n^{-p} \quad \forall n \in \mathbb{N}^+. \tag{4.1}$$

Thus, using the Definition 14 one can also state that the convergence error $|e_n| = |x_n - L|$ is $\mathcal{O}(n^{-p})$. When addressing the convergence performance in tests, the convergence of order p will be given by the following expression:

$$p = \frac{\log \frac{|e_{n-1}|}{|e_n|}}{\log \frac{2^n}{2^{n-1}}}. \tag{4.2}$$

If $p = 2$ we say that the order of the convergence is *quadratic*, i.e., if we double the grid size the error will drop by a factor of 4. In the work of Lord et al. (2008) the error of convergence was defined as being the ratio between two consecutive errors.

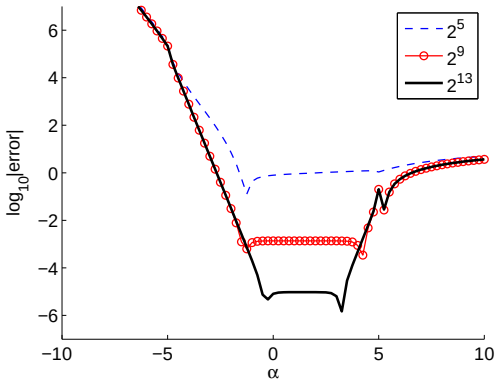
4.2 Sensitivity to the integration parameters

In the next two subsections the effects of the choices for the damping factor (α), and the integration range (L) are studied. For each case, a value will be chosen and used throughout the remaining tests of this chapter. The reference values will be calculated using the Carr-Madan method with a grid size of 2^{20} points.

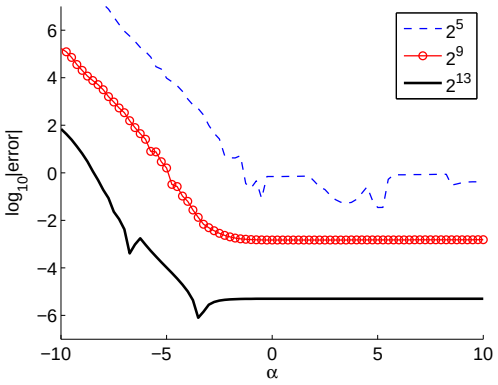
[1] This method was also used as a reference value for the European-style option tests in Lord et al. (2008).

4.2.1 Damping factor α

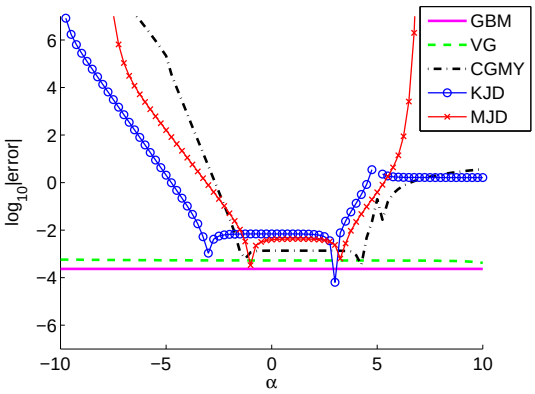
Considering the Carr-Madan method for European-style options, Lord & Kahl (2007) have demonstrated how to approximate the optimal damping coefficient when the payoff transform is known, which increases the numerical stability of the method. Their study shows that using the optimal value is particularly effective for *in/out-of-the-money* options and options with short maturities. Nevertheless, their rationale is more complicated to carrie over to the Bermudan-style options, since at each exercise opportunity the European option will have a different level of *moneyness* and therefore a different value of α will be needed each exercise opportunity. Thus, which single choice for α will be optimal is a complex problem. The approach in this work was to test a range of α's under different grid sizes, asset price processes and option types. In Figure 4.1 we show the error of the CONV method as a function of α. Figure 4.1(a) shows a European put under T3-CGMY, and in Figure 4.1(b), a Bermudan put under T2-VG is displayed. All tests in Figure 4.1 were performed considering $S(t_0) = 100$, $K = 110$ and $T = 0.1$, where T represents the time to maturity.



(a) European put for T3 - CGMY and T=0.1.



(b) Bermudan put for T2 - VG and T=1.



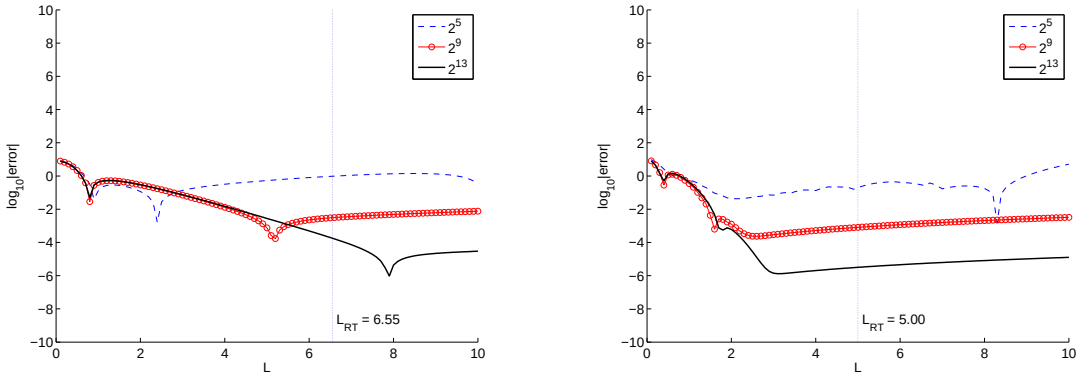
(c) European put under all processes being studied.

Figure 4.1: Error of the CONV method as a function of α.

From Figure 4.1 it is possible to conclude that, there exists a wide range of values of α for which the error is small and stable. To further confirm this, we present in Figure 4.1(c) a test for a European put using a grid size with 2^9 points and under all asset price processes being studied. Thus, we will set $\alpha = 0$ for all further experiments, which produces good results for the studied examples.

4.2.2 Integration range L

As mentioned in Section 3.4, the value of the integration range L cannot be too small to avoid large truncation errors, but cannot also be large enough, such that, will create errors on the Fourier domain. In Figure 4.2 we present the error of the CONV method as a function of L and also mark the value of L calculated using the rule of thumb of equation (3.33). Regarding the choice of δ in equation (3.33), we follow Lord et al. (2008) by setting $\delta = 20$ under GBM and for the other asset price models, which have fatter tails, we use $\delta = 40$. Figure 4.2(a) shows the results for an European option under T6-KJD and with $T = 0.5$. In Figure 4.2(b), we present the results for a Bermudan option under T1-GBM using $T = 1$ and $M = 10$.



(a) European call for T6-KJD and $T = 0.5$.

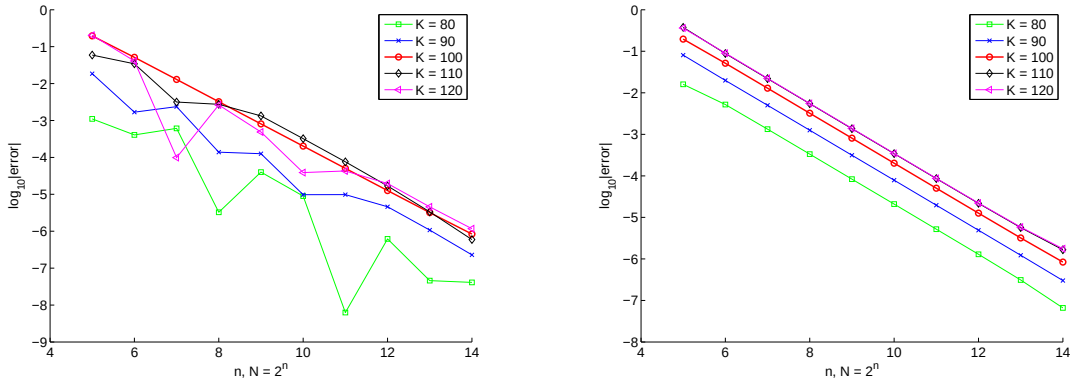
(b) Bermudan call for T1-GBM and $T = 1$.

Figure 4.2: Error of the CONV method as a function of L.

The results from Figure 4.2 show that the CONV method is stable for a wide range of values of L and that the error increases for smaller and larger values of L, as expected. We can further conclude that the calculated value of L using the rule of thumb is not optimal but produces satisfactory results for both examples. Hence, for the remaining tests of this chapter we will use the rule of thumb formula (3.33).

4.3 European options

In this section, the CONV method for pricing European options under several asset price processes is evaluated. The reference value will be calculated again using the Carr-Madan formula (with $N = 2^{20}$ points), except when mentioned otherwise. The first tests check the accuracy of the two discretization methods presented in Section 3.4, Discretization I and II. In Figure 4.3, we plot the convergence error of the two discretization methods for pricing a European call option at various strikes under T2-VG.



(a) Discretization I.

(b) Discretization II.

Figure 4.3: Convergence of the two discretization methods for pricing European call options at several K values under T2-VG.

From Figure 4.3(a) we notice that the only option with a stable convergence in Discretization I is the *at-the-money* option with $K = 100$. Comparing with Figure 4.3(b), it is clear that placing the strike on the y-grid in Discretization II ensures a stable convergence for any degree of *moneyness*. Therefore, for the remaining tests of this chapter the Discretization II will be used. Next, in Table 4.1, the accuracy, speed and convergence order of the CONV algorithm for pricing European call options is shown.

Table 4.1: Speed, error and convergence order for European call options under T1-GBM and T2-VG, with $K = 90$ and $T = 0.1$.

$(N = 2^n)$	GBM: $V_{ref}(t_0, S(t_0)) = 11.1352431$			VG: $V_{ref}(t_0, S(t_0)) = 10.9937031$		
n	Time (ms)	Error	p	Time (ms)	Error	p
7	0.73	$-2.08e-3$	-	0.61	$-4.49e-5$	-
8	0.67	$-5.22e-4$	2.0	0.66	$-1.18e-5$	1.9
9	0.71	$-1.30e-4$	2.0	0.76	$1.48e-6$	3.0
10	0.92	$-3.26e-5$	2.0	0.97	$7.75e-6$	-2.4
11	1.32	$-8.15e-6$	2.0	1.59	$1.23e-5$	-0.7
12	2.02	$-2.04e-6$	2.0	2.56	$1.32e-5$	-0.1

From Table 4.1 we can observe that the convergence under GBM is clearly of regular second order. Regarding the VG model, one can note that, we are in the presence of nonsmooth convergence, yet with

an accuracy superior to the GBM case for smaller grid sizes. In Lord et al. (2008, p. 1696), the nonsmooth convergence of the second case is explained by the presence of an "highly oscillatory integrand".

Next, we performed the evaluation of each model using the following pricing methods: BM, CM and CONV. These tests examine if any method is superior to the others, and also compares the performance of the CONV against the remaining methods. The results for these tests are summarized in Table 4.2, where CF stands for the option valuation using the closed-form solution. Thus, the reference value for these tests was, naturally, the value from the closed-form solution which is present in the column labeled as $V_{ref}(t_0, S(t_0))$. The tests were performed using the following parameters: $S(t_0) = 100$, $K = 110$, $r = 0.1$, $q = 0$, $T = 0.1$ and the option was an European put. For the CGMY model the T3-CGMY was the selected test case.

Table 4.2: Speed and accuracy analysis of the CM, CM and CONV pricing methods under T1-GBM, T2-VG, T3-CGMY, T5-KJD and T6-MJD.

Mths	CF		BM		CM		CONV	
Mdls	Time	$V(t_0, S(t_0))$	Time	Error	Time	Error	Time	Error
GBM	2	9.4950978	15	0	5	$5.60e-1$	1	$-2.34e-4$
VG	8	9.1438446	1068	$-2.10e-1$	4	$-1.17e-1$	3	$-2.11e-1$
CGMY	1255	10.6692757	11	$-2.15e-7$	2	$3.32e-1$	1	$-1.37e-3$
KJD	429	9.2609414	6	$-3.85e-12$	2	$4.07e-1$	1	$-6.95e-3$
MJD	27	9.1953362	6	$1.06e-10$	1	$4.50e-1$	1	$-3.96e-3$

Analyzing Table 4.2 one draw several conclusions. First, observing the reference values of the closed-form solutions, one can see that the values do not differ much from one model to the others. Note that the solution for the CGMY is not from a closed-form expression, but from the Carr-Madan formula, as mentioned in Section 2.8 (using 2^{20} grid points). This explains why the running time of the closed-form solution is much greater in this case. Examining all errors, one can infer that the VG process was the most difficult case to price, since the errors for this case are higher when compared with the other models. Even for the BM, which has negligible errors for the other models, the VG has a significant error. In terms, of accuracy, the BM method beats the competition, but one most note that the CM and CONV methods were priced using a grid size wit 2^9 points; using higher grid sizes for these methods would yield in smaller errors. In terms of speed, the CONV method is, definitely, the fastest one: it offers the lowest running times for every model. The CONV method is not only fast but also extremely accurate. The pricing values can agree up to three decimal digits using a small grid size with 512 ($= 2^9$) points. As a final remark, we can conclude that accuracy has a price in terms of speed, even the closed-form solutions are hard to calculate (specially for the KJD and VG). Therefore, the best solution for obtaining fast and accurate prices is the CONV method.

The next test performs a speed and accuracy analysis for a wide range of grid sizes, with $n \in [7, 18]$ ($N = 2^n$). For this test, the CPU running times were only averaged after 100 experiments. Figure 4.4

shows the errors of the CM and CONV methods as a function of the running times (which are proportional to the grid sizes), for pricing a European call under T6-MJD with $K = 80$ and $T = 0.5$.

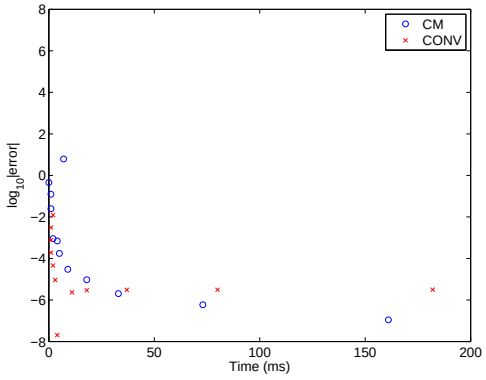


Figure 4.4: Speed and accuracy of the CM and CONV methods for pricing a European call under T6-MJD with $K = 80$ and $T = 0.5$.

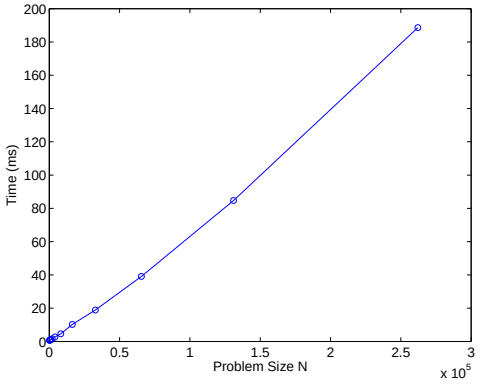


Figure 4.5: Running time of the CONV method as a function of the grid size N.

Both methods present similar results and we can conclude that no method is superior to the other for all grid sizes. Even so, for small grid sizes the CONV method displays a higher performance, being faster and accurate. As for large grid sizes, the situation is the reverse, being the CM method the one with superior performance. In Figure 4.4, we can also observe an experiment for the CONV method with a small grid size and the lowest error. This "outlier" may be owed to that specific grid size having an optimal α, which made the error so small. Finally, for concluding the study of the CONV method using European-style options we show, in Figure 4.5, the running time as a function of the grid size. The test scenario is the same of the previous example (T6-MJD). It is evident from Figure 4.5 that the order of growth or complexity of the method is slightly worst than *linear*, most likely *linearithmic* or even *logarithmic*. This was expected, since the complexity of the method for European-style options is the same as the FFT, i.e., $\mathcal{O}(N \log N)$.

4.4 Bermudan options

In this section, the pricing of Bermudan-style options will be evaluated in terms of speed, accuracy and order of convergence. Table 4.3 summarizes the results of the tests performed for a 10-times exercisable Bermudan put under T1-GBM and T2-VG, with $K = 90$ and $T = 0.1$. In these tests, the reference value was obtained using the CONV method with 2^{20} grid points, and the CPU times were only averaged after 100 experiments.

From the results in Table 4.3 we can observe that the GBM maintains a regular second order convergence. In spite of displaying a nonsmooth convergence for the VG case, the CONV method yields a similar accuracy for both models. The stable convergence of the algorithm enables the use of an ex-

Table 4.3: Speed, error and convergence order for a 10-times exercisable Bermudan put under T1-GBM and T2-VG, with $K = 90$ and $T = 0.1$.

$(N = 2^n)$	GBM: $V_{ref}(t_0, S(t_0)) = 11.98745352$			VG: $V_{ref}(t_0, S(t_0)) = 9.04064612$		
n	Time (ms)	Error	p	Time (ms)	Error	p
7	9.25	$-2.70e-2$	-	10.03	$-3.64e-2$	-
8	9.99	$-7.04e-3$	1.9	10.51	$-1.35e-3$	4.8
9	11.07	$-1.90e-3$	1.9	12.29	$-1.49e-3$	-0.1
10	14.10	$-4.96e-4$	1.9	16.29	$-2.59e-4$	2.5
11	20.92	$-1.25e-4$	2.0	26.44	$-7.37e-5$	1.8
12	32.63	$-3.02e-5$	2.0	42.70	$-2.12e-5$	1.8

trapolation procedure. The running times of the CONV algorithm were increased by a factor of 10, approximately, when compared to the ones for European-style options — see Table 4.1. This was expected given that, in Section 3.5, we determined an order of growth of $\mathcal{O}(MN \log N)$ for the Bermudan pricing algorithm, and we have a 10-times exercisable Bermudan option. It should be noted that, in Lord et al. (2008) work, the increase of the running time of the algorithm, was not so significant. In Lord et al. (2008) the running times were only increased by a factor between two and four.

4.5 American options

Finally, we conclude the numerical experiments of the CONV method with the pricing of American-style options. Given that this type of option can be exercisable at any time, the number of exercise opportunities is much greater, and thus, the running time of the algorithm will be increased. Hence, the CPU times were only averaged after 10 experiments. Furthermore, the reference value was obtained on a grid with 2^{14} points and using a 2-times repeated Richardson extrapolation[2] on 512-, 256- and 128-times exercisable Bermudan options.

First, we compare the accuracy and CPU time of the two approximation methods referred in Section 3.6, that is, the direct approximation via a Bermudan option with many exercise opportunities and the repeated Richardson extrapolation technique. For the former we will use a Bermudan option with $N/2$ exercise opportunities, i.e., if we have a grid size with $N = 2^n$ points the number of exercise times will be $M = 2^{n-1}$. In the latter case, we opted for 2 extrapolations on 3 Bermudan options with 128, 64, and 32 exercise opportunities, which yielded stable and satisfactory results. In this first test, we price an American put under T1-GBM with $K = 90$ and $T = 0.1$. The performance of both approaches is summarized in Table 4.4, where $P(N/2)$ denotes the approximation using an $N/2$-times exercisable Bermudan option, while "Richardson" denotes the results obtained by the 2-times repeated Richardson extrapolation scheme.

[2]In all experiments we follow Lord et al. (2008) by considering $\gamma = 1$ in equation (3.39).

Table 4.4: Comparison of two approximation methods in terms of speed, error and convergence order for an American put under T1-GBM, with $K = 90$, $T = 0.1$ and $V_{ref}(t_0, S(t_0)) = 12.16941552$.

$(N = 2^n)$	$P(N/2)$			Richardson		
n	Time (ms)	Error	p	Time (ms)	Error	p
7	58	$-5.75e-2$	-	204	$-2.31e-2$	-
8	124	$-2.23e-2$	1.4	217	$-7.85e-3$	1.6
9	277	$-9.31e-3$	1.3	243	$-2.06e-3$	1.9
10	679	$-4.15e-3$	1.2	303	$-5.25e-4$	2.0
11	2064	$-1.94e-3$	1.1	465	$-1.27e-4$	2.0
12	6431	$-9.37e-4$	1.1	732	$-1.94e-5$	2.7

Analyzing the previous results, it is evident that the extrapolation-based method converges faster and costs far less CPU time than the direct approximation approach. Only for small grid sizes does the direct method run faster, which is expected since it only prices one option. Another aspect to note is that the method with extrapolation scheme always displays an higher accuracy than the direct method. Comparing our results with the ones from Lord et al. (2008), we can observe that the running times here are greater than the ones reported in the article. However, this was expected given that the calculations are dependent on the performance of the Bermudan pricing method. Furthermore, the extrapolation mechanism in this work displays faster results than in Lord et al. (2008). This can be inferred, by noting that, in Lord et al. (2008) tests the running times double with each increase in the grid size, yielding a *linear* order of growth for this component. Yet, in our results, the order of growth is *sub-linear*, given that, to obtain the double of the running time for the case with $n = 7$, the grid size was increased by a factor of 16 $(= 2^4)$ times.

Having confirmed that the repeated Richardson extrapolation yields faster and accurate results, we turn our attention to the pricing of American options under alternative dynamics, by using the VG and both CGMY test sets. The second CGMY test case, T4-CGMY, has $Y > 1$ and is considered a hard test case when numerically solving the corresponding PIDE — see Almendral (2007). In Table 4.5 we show the results for American put options under VG and CGMY, in terms of speed and accuracy.

Table 4.5: Speed and errors for American put options under VG and CGMY.

	T2-VG		T3-CGMY		T4-CGMY	
	$K = 110, T = 1$		$K = 1, T = 1$		$K = 98, T = 0.25$	
$(N = 2^n)$	$V_{ref} = 9.99944573$		$V_{ref} = 0.11215935$		$V_{ref} = 9.22543580$	
n	Time (ms)	Error	Time (ms)	Error	Time (ms)	Error
7	225	$6.94e-2$	285	$4.93e-4$	284	$9.71e-2$
8	242	$-2.91e-2$	314	$2.36e-4$	315	$2.61e-2$
9	282	$-7.24e-3$	374	$-6.12e-5$	381	$-5.53e-4$
10	364	$-2.72e-4$	495	$3.16e-5$	519	$-1.48e-5$
11	604	$-2.05e-3$	800	$-6.12e-6$	849	$-1.63e-4$
12	926	$-5.59e-3$	1464	$-2.78e-6$	1563	$-7.62e-5$

These parameter sets stem from the PIDE literature — see Table A.1 — where it were reported the following values for the option prices: 10 for the T2-VG, 0.112171 for T3-CGMY and 9.2254842 for T4-CGMY. The reference values for both CGMY cases agree up to four digits with the values from the literature, while the VG test case only agrees up to three digits. Observing Table 4.5, we can conclude that the results are good, which proves that the CONV method is able to price American options under a wide variety of Lévy processes.

5 Conclusions and Future work

I am turned into a sort of machine for
observing facts and grinding out conclusions.

Charles Darwin, Author of the theory of
evolution by natural selection.

In this final chapter, we conclude this thesis by outlining our main findings and contributions in Section 5.1, and by discussing directions for future work in Section 5.2.

5.1 Conclusions

In this thesis we have presented a recent FFT-based method, based on the work of Lord et al. (2008), for pricing options with early-exercise features, dubbed as the CONV method. In the same way as other methods, based on the knowledge of the characteristic function, it is flexible with respect to the choice of asset price process and the type of option contract. This feature has been demonstrated in numerical examples for European, Bermudan, and American-style options and for five Lévy driven processes, namely, GBM, VG, CGMY, MJD and KJD. The key assumption of the method is that the underlying assets are driven by processes with independent increments, whose characteristic function is readily available. The main strengths of the method are its flexibility, computational speed and accuracy. It was shown how the usage of the FFT enabled us to achieve a complexity of $\mathcal{O}(MN \log N)$, where N is the number of grid points and M is the number of exercise opportunities of the option contract. This constitutes an improvement in comparison with the QUAD method of Andricopoulos et al. (2003), which has a complexity of $\mathcal{O}(MN^2)$.

Regarding the numerical experiments we were able to come to several conclusions. First, we evaluated the effect of two of the CONV's method parameters, the damping factor and size of integration range, where we observed that there were a wide range of values for which the method performed accurately. Next, the importance of avoiding discontinuities during the integration was emphasized, by comparing two discretization procedures, where the one with attention to discontinuities achieved a smoother convergence. Still during the European options tests, we were able to compare the speed and accuracy of the CONV method with two standard methods of the literature. The results showed

that, the CONV method was the fastest, and yielded a fine accuracy even when the grid size was small. Finally, we completed the European option tests, by comparing the speed and accuracy of the CONV mehtod with the Carr-Madan formula, showing that if one has no time and memory constraints, the Carr-Madan method has an higher accuracy, otherwise the CONV method is the most efficient choice. The tests on the early-exercise options, starting with the Bermudan options, demonstrated that the speed and accuracy of the method was carried over to the cases with early-exercise features. The obtained running times were above the ones mentioned in Lord et al. (2008), but in our point of view, are in line with what was expected from theory. Finally, our study showed that it was possible to price American options quite quickly, and with an acceptable level of accuracy. The superiority of the repeated Richardson extrapolation was shown to be evident, unless one is restricted for using a very small grid. The prices from the CONV method were also compared to values from the PIDE literature, and shown to agree up to three or four decimal digits.

The conclusion of this experiment is that the CONV method is a flexible and powerful tool, in the context of exponential Lévy driven models. Given its speed and flexibility, it has an edge over PIDE schemes for the pricing of options with early-exercise features, in particular for exponential Lévy models with infinite activity. Nevertheless, a more complete set of tests would have to be performed. It is worth noting that there will always be special cases where an highly efficient PIDE scheme can be designed and tunned for a given model, but where the flexibility of pricing several models is lost. The speed of the method can make it possible to calibrate models to the prices of American options, which are the most exchange traded option contracts.

5.2 Future work

In this section, we conclude this thesis by suggesting further research directions, which we consider more obvious or most relevant, in context of the CONV method, for future research endeavors.

Optimal alpha for early-exercise options As discussed in Section 4.2, the problem of choosing the optimal damping factor α is a complex problem to solve when considering options with early-exercise features. Obtaining a solution for this problem with low computational costs would improve the accuracy of pricing method. It would also facilitate the application of the CONV method, not requiring a fine tune of the parameter when the maximum accuracy is required.

Interpolation issues In Section 3.4, we presented a simple linear interpolation algorithm for determining the location of discontinuities. Often this solution is regarded as an adequate algorithm that yields sufficiently accurate results. Nevertheless, McCulloch (2003) argues that simple linear interpolation may give an interpolation error in excess of the Fourier inversion computational error due to the convexity of the pricing function. The usage of cubic spline interpolation could provide a capable way to capture

the curvature of these pricing functions, and therefore increase the accuracy and stability of the method.

Gains with FrFT. During the discretization procedure one is restricted, by the Nyquist relation, to have the log-spot price grid sparse and the Fourier grid coarse or *vice-versa*. To overcome this constraint, a generalization of the conventional discrete Fourier transform was created, and named as the Fractional Fourier Transform (FrFT). The benefit of employing a FrFT instead of running a single FFT is due to the fact that a FrFT with smaller N may achieve the same accuracy as using a FFT with much larger N. In Lord et al. (2008, p. 1685) it is referred that the numerical tests indicated that the advantages of the FrFT do not outweigh the speed of the FFT. Even so, a more detailed study on this topic would be highly valuable.

Efficient pricing of multi-asset options The extension of the CONV method to multiple dimensions has been studied in Leentvaar & Oosterlee (2008), where it was shown that the method is able to efficiently price multi-asset options of European and early-exercise type under Lévy price dynamics. To accomplish this feat, a partitioning technique was employed, which enables one to speed up the computation, by distributing some multi-dimensional splitted parts over a system of parallel computers. Hence, future research should be directed to the evaluation of parallel methods on machines with a significant number of processors.

A Resources

A.1 Parameters sets

This section presents Table A.1 with the six parameter sets which are used in this work. The last column has references for articles where the parameters are also used. The first four parameter sets were also used in Lord et al. (2008).

Table A.1: Parameter sets in numerical experiments.

Test ID	$S(t_0)$	r	q	Model parameters			References
T1-GBM	100	0.1	0	$\sigma = 0.25$			
T2-VG	100	0.1	0	$\sigma = 0.12$	$\theta = -0.14$	$\nu = 0.2$	Madan et al. (1998) O'Sullivan (2005) Maller et al. (2006)
T3-CGMY	1	0.1	0	$C = 1$ $Y = 0.5$	$G = 5$	$M = 5$	Almendral (2007)
T4-CGMY	90	0.06	0	$C = 0.42$ $Y = 1.0102$	$G = 4.37$	$M = 191.2$	Wang et al. (2007)
T5-KJD	100	0.05	0	$\sigma = 0.16$ $\eta_1 = 10$	$p = 0.4$ $\eta_2 = 5$	$\lambda = 1$	Kou (2002)
T6-MJD	100	0.0075	0	$\sigma = 0.2$ $\sigma_Y = 0.6$	$\lambda = 0.01$	$\mu_Y = -0.2$	

A.2 Exponential distribution

Definition 16 (Exponential random variable) *A positive random variable X follows an exponential distribution with parameter $\lambda > 0$ if it has probability density function given by*

$$f(x) = \lambda e^{-\lambda x} \mathbf{1}_{\{x \geq 0\}}, \tag{A.1}$$

with the distribution function, defined for $y \in [0, \infty[$, with the following expression

$$F_X(x) = 1 - e^{-\lambda x}. \tag{A.2}$$

Definition 17 (Absence of memory) *Consider the $X \geq 0$ random variable such that*

$$\forall t, s > 0, \quad \mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s). \tag{A.3}$$

Then X has an exponential distribution.

A.3 Gamma process

This section describes the gamma process which is also one of the basic Lévy processes that receives its name due to its increments having a gamma distribution. The formal definition is given next:

Definition 18 (Gamma process) *A stochastic process $\gamma = \{\gamma_t, t \geq 0\}$ is a gamma process if it satisfies all conditions of a Lévy process and the increment $\gamma_{t+s} - \gamma_t$ has a gamma distribution with parameters $c > 0$ and $\lambda > 0$, for $s > 0$. The density of the increment is given by:*

$$f_\gamma(x) = \frac{\lambda^c}{\Gamma(c)} x^{c-1} e^{-\lambda x} \qquad x > 0,$$

where Γ represents the gamma function that can be calculated for a positive real number t as:

$$\Gamma(t) = \int_0^\infty \frac{x^t e^{-x}}{x} dx.$$

This is a pure jump increasing Lévy process with the Lévy measure given by

$$\nu(x) = \frac{c e^{-\lambda x}}{x} \mathbf{1}_{\{x>0\}},$$

where λ controls the decay rate of the big jump arrivals and c is the scaling parameter that changes the size of all jumps simultaneously, basically allowing the change of time scale. Often this process is parameterized by specifying the mean μ and variance v that one desires the process to have. Hence, next we present expressions for obtaining the process parameters in terms of the given distribution properties:

$$c = \frac{\mu^2}{v} \qquad \lambda = \frac{\mu}{v}.$$

Thus the characterisitc function associated with the gamma density can be defined in terms of μ and v as

$$\phi_\gamma(u) = \left(\frac{1}{1 - iu\frac{v}{\mu}} \right)^{\frac{\mu^2 t}{v}}.$$

Bibliography

Almendral, A. (2007). Accurate evaluation of European and American options under the CGMY process. *SIAM J. Sci. Comput.*(29), 93–117.

Andricopoulos, A. D., Widdicks, M., Dmitrasinovic-Vidovic, M. A. H., & Newton, D. P. (2003). Universal option valuation using quadrature methods. *Journal of Financial Economics*, *67*(3), 447–471.

Bachelier, L. (1900). *Théorie de la spéculation*. Unpublished doctoral dissertation, Annales scientifiques de l'École Normale Supérieure.

Bakshi, G., & Madan, D. (2000). Spanning and derivative-security valuation. *Journal of Financial Economics*(55), 205–238.

Barndorff-Nielsen, O. E. (1995). *Normal inverse Gaussian distributions and the modeling of stock returns* (Tech. Rep. No. 300). Department of Theoretical Statistics, Aarhus University.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society Series B*, *63*(2), 167-241.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *64*(2), 253–280.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *Review of Financial Studies*, *9*(1), 69-107.

Bernstein, P. (1996). *Against the gods: The remarkable story of risk*. New York [u.a.]: Wiley.

Black, F., & Scholes, M. S. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637-54.

Boyle, P. P. (1977). Options: A Monte Carlo approach. *Journal of Financial Economics*, *4*(3), 323-338.

Carr, P., Geman, H., Madan, D., & Yor, M. (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business*, *75*(2), 305–332.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for Lévy processes. *Mathematical Finance*, *13*(3), 345-382.

Carr, P., & Madan, D. B. (1999). Option valuation using the fast Fourier transform. *Journal of Computational Finance*, *2*, 61–73.

Chang, C.-C., Chung, S.-L., & Stapleton, R. C. (2007). Richardson extrapolation techniques for the pricing of American-style options. *Journal of Futures Markets*, *27*(8), 791–817.

Cont, R., & Tankov, P. (2012). *Financial modelling with jump processes*. Chapman & Hall/CRC.

Courtault, J.-M., Kabanov, Y., Bru, B., Crépel, P., Lebon, I., & Le Marchand, A. (2000). Louis Bachelier on the centenary of théorie de la spéculation. *Mathematical Finance*, *10*(3), 339–353.

Cox, J. (1975). *Notes on option pricing I: Constant elasticity of diffusions* (Tech. Rep.). Stanford University.

Cox, J., & Ross, S. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, *3*, 145–166.

Cox, J., Ross, S., & Rubenstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, *7*, 229–263.

Davis, M., Bachelier, L., Etheridge, A., & Samuelson, P. (2011). *Louis Bachelier's theory of speculation: The origins of modern finance*. Princeton, New Jersey: Princeton University Press.

Eberlein, E., Keller, U., & Prause, K. (1998). New insights into smile, mispricing, and value at risk: The hyperbolic model. *Journal of Business*, *71*(3), 371-405.

Eberlein, E., & Prause, K. (1998). *The generalized hyperbolic model: Financial derivatives and risk measures.* Springer.

Fang, F., & Oosterlee, K. (2008). *Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions* (MPRA Paper No. 9248). Germany: University Library of Munich.

Geman, H. (2002). Pure jump Lévy processes for asset price modeling. *Journal of Banking and Finance*, *26*(7), 1297–1316.

Geske, R., & Johnson, H. E. (1984). The American put option valued analytically. *Journal of Finance*, *39*(5), 1511-24.

Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika*(37), 481–482.

Gurland, J. (1948). Inversion formulae for the distribution of ratios. *Annals of Mathematical Statistics*(19), 228–237.

Harrison, J. M., & Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, *20*(3), 381-408.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, *6*, 327–343.

Hull, J., C. (2009). *Options, futures, and other derivatives* (7 ed.). Upper Saddle River, New Jersey: Prentice Hall.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Manage. Sci.*, *48*(8), 1086–1101.

Leentvaar, C., & Oosterlee, C. (2008). Multi-asset option pricing using a parallel Fourier-based technique. *Journal of Computational Finance*, *12*(1), 1–26.

Lewis, A. L. (2001). *A simple option formula for general jump-diffusion and other exponential Lévy processes* (Related articles). Finance Press.

Lord, R., Fang, F., Bervoets, F., & Oosterlee, C. W. (2008). A fast and accurate fft-based method for pricing early-exercise options under Lévy processes. *SIAM J. Sci. Comput.*, *30*(4), 1678–1705.

Lord, R., & Kahl, C. (2007). Optimal Fourier inversion in semi-analytical option pricing. *Journal of Computational Finance*, *10*(4), 1–30.

Madan, D. B., Carr, P., & Chang, E. C. (1998). The variance gamma process and option pricing. *European Finance Review*, *2*, 79–105.

Madan, D. B., & Milne, F. (1991). Option pricing with V.G. martingale components. *Mathematical Finance*, *1*, 39–55.

Madan, D. B., & Seneta, E. (1990). The variance gamma (v.g.) model for share market returns. *Journal of Business*, *63*(4), 511-24.

Maller, R., Solomon, D., & Szimayer, A. (2006). A multinomial approximation for American option prices in Lévy process models. *Mathematical Finance*, *16*(4), 613–633.

Mandelbrot, B., & Hudson, R. (2010). *The (mis)behaviour of markets: A fractal view of risk, ruin and reward*. Pine Street, London: Profile.

Matsuda, K. (2004). *Introduction to Merton jump diffusion model* (Tech. Rep.). The Graduate Center, The City University of New York.

McCulloch, J. H. (2003). *The risk-neutral measure and option pricing under log-stable uncertainty* (Working Papers No. 03-07). Ohio State University, Department of Economics.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics*, *4*(1), 141-183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1-2), 125-144.

Merton, R. C. (1997). *Applications of option-pricing theory: Twenty-five years later.* Nobel Lecture.

Nocedal, J., & Wright, S. J. (2000). *Numerical optimization*. Springer.

Options Exchange, C. B. of. (2013). *Cboe history.* `http://www.cboe.com/aboutcboe/History.aspx`.

O'Sullivan, C. (2005). *Path dependant option pricing under Lévy processes.* EFA 2005 Moscow Meetings Paper.

Papapantoleon, A. (2008). An introduction to Lévy processes with applications in finance. *ArXiv e-prints*.

Pliska, S. R. (2010). *A history of options from the middle ages to Harrison and Kreps.* Presentation at the Fields Institute, Toronto, Canada.

Press, S. (1967). A compound events model for security prices. *Journal of Business*(40), 317–335.

Raible, S. (2000). *Lévy processes in finance: Theory, numerics, and empirical facts*. Unpublished doctoral dissertation, Albert-Ludwigs, Universitat Freiburg.

Ramezani, Y. Z., C. A. (1999). *Maximum likelihood estimation of asymmetric jump-diffusion process: Application to security prices.* (Working paper)

Reiner, E. (2001). *Convolution methods for path-dependent options.* Presentation at Financial Mathematics workshop.

Samuelson, P. A. (1965). Rational theory of warrant pricing. *Industrial Management Review*, *6*, 13–31.

Schoutens, W. (2001). *Meixner processes in finance* (Tech. Rep. Nos. 2002–004). EURANDOM, Eindhoven.

Schoutens, W. (2003). *Lévy processes in finance: Pricing financial derivatives* (1 ed.). Wiley.

Schwartz, E. S. (1977). The valuation of warrants: Implementing a new approach. *Journal of Financial Economics*, *4*(1), 79-93.

Sedgewick, R., & Wayne, K. (2011). *Algorithms* (4 ed.; Addison-Wesley, Ed.). Pearson Education.

Wang, I., Wan, J., & Forsyth, P. (2007). Robust numerical valuation of European and American options under the cgmy process. *Journal of Computational Finance*, *10*(4), 31–69.